On approximations, complexity, and applications for copositive programming

Gijben, Luuk

On Approximations, Complexity, and Applications for Copositive Programming

Luuk Gijben

This PhD project was carried out at the Johann Bernoulli Institute for Mathematics and Computer Science at the University of Groningen according to the requirements of the Groningen Graduate School of Science.

Funding was provided by the Netherlands Organisation for Scientific Research (NWO) through Vici grant no.639.033.907.

rijksuniversiteit
groningen

On Approximations, Complexity, and Applications for Copositive Programming

Proefschrift

ter verkrijging van de graad van doctor aan de
Rijksuniversiteit Groningen
op gezag van de
rector magnificus prof. dr. E. Sterken
en volgens besluit van het College voor Promoties.

De openbare verdediging zal plaatsvinden op

vrijdag 23 januari 2015 om 16.15 uur

door

Luuk Gijben
geboren op 30 november 1986
te Boxtel

Promotor
Prof. dr. M. Dür

Beoordelingscommissie
Prof. dr. E. de Klerk
Prof. dr. F. Rendl
Prof. dr. H.L. Trentelman

ISBN: 978-90-367-7527-4 (printed version)
ISBN: 978-90-367-7526-7 (electronic version)

printed via www.proefschrift-aio.nl

Acknowledgements

I have found that my time in Groningen as a PhD-student has been a very interesting experience that has given me a lot of wonderful opportunities. This experience has been made possible by a number of people to whom I owe a great deal of gratitude. I would like to take this opportunity to thank all of you. Furthermore there are several people I would like to thank separately for their support.

To start of I would like to thank my supervisor, Prof. Dr. M. Dür for giving me the opportunity to pursue a PhD at the University of Groningen. Moreover I would like to thank the University of Groningen, The Johann Bernoulli Institute and the Groningen Graduate School for their assistance and the resources they offered.

Furthermore I have had the opportunity to meet and work with a number of people in the past four years. In particular I would like to thank Dr. Peter Dickinson for a number of fruitful discussions resulting in several successful publications. By the same token I would also like to express my gratitude to Dr. Roland Hildebrand. I would furthermore like to thank Dr. Juan Vera for a series of very interesting and eye opening conversation that have given me a better understanding of how to approach doing research.

During my time in Groningen I have also received a lot of support from my friends and family. For this I am grateful, particularly I would like to thank my parents and my brother who are always there for me.

Finally I would like to thank Michi for being part of my life and for being the person she is. She has brought color to my life in so many ways for which I will be eternally grateful.

Contents

Personal contributions

During my time as a PhD-student at the University of Groningen I have produced along with my co-authors the following papers.

Published articles:

[DG14] Peter J.C. Dickinson and Luuk Gijben. On the computational complexity of membership problems for the completely positive cone and its dual. *Computational Optimization and Applications*, 57(2):403-415, 2014.

[DDGH13a] Peter J.C. Dickinson, Mirjam Dür, Luuk Gijben, and Roland Hildebrand. Irreducible elements of the copositive cone. *Linear Algebra and its Applications*, 439(6):1605-1626, 2013.

[DDGH13b] Peter J.C. Dickinson, Mirjam Dür, Luuk Gijben and Roland Hildebrand. Scaling relationship between the copositive cone and Parrilo's first level approximation. *Optimization Letters*, 7(8):1669-1679, 2013. *(won the 2013 OPTL Best Paper Award)*

Articles in construction:

Mirjam Dür and Luuk Gijben. Graph Isomorphism and Copositive Programming.

Besides these papers I have obtained several separate currently unpublished results. What follows is a more detailed explanation of my personal contributions to the field of copositive programming as presented in this thesis.

It has been shown by Murty and Kabadi [MK87] that the strong membership problem for the copositive cone, that is deciding whether or not a given matrix is in the copositive cone, is a co-NP-complete problem. From this it has long been assumed that this implies that the question of whether or not the strong membership problem for the dual of the copositive cone, the completely positive cone, is also an NP-hard problem. However, the technical details for this have not previously been looked at to confirm that this is indeed the case. In *Chapter* 2, which is based on [DG14], I prove together with Dr. Peter J.C. Dickinson that the strong membership problem for the completely positive cone is indeed NP-hard. Furthermore, we show that even the so called weak membership problems for both of these cones are NP-hard. An alternative proof of the NP-hardness of the strong membership problem for the copositive cone is also provided.

Chapter 3 comes from the paper [DDGH13a] I published together with Dr. Peter J.C. Dickinson, Prof. Dr. Mirjam Dür and Dr. Roland Hildebrand. An element A of the $n \times n$ copositive cone $\mathcal{COP}^n$ is defined to be *irreducible* with respect to the nonnegative cone if it cannot be written as a nontrivial sum $A = C + N$ of a copositive matrix C and an element-wise nonnegative matrix N (note that this concept of irreducibility differs from the standard one normally studied in matrix theory). This property was studied by Baumert [Bau65] who gave a characterization of irreducible matrices. It is demonstrated in this chapter that Baumert's characterization is incorrect and a correct version of his theorem is given instead. This establishes a necessary and sufficient condition for a copositive matrix to be irreducible. For the case of 5×5 copositive matrices a complete characterization of all irreducible matrices is given. It is shown that those irreducible matrices in $\mathcal{COP}^5$ which are not positive semidefinite can be parameterized in a semi-trigonometric way. Finally, a proof is given for the result that every 5×5 copositive matrix which is not the sum of a nonnegative and a positive semidefinite matrix can be expressed as the sum of a nonnegative and a single irreducible matrix.

This result is then used in *Chapter* 4, the content of which is from the paper [DDGH13b] (which received the 2013 OPTL Best Paper Award) that I also published with Dr. Peter J.C. Dickinson, Prof. Dr. Mirjam Dür and Dr. Roland Hildebrand. In particular we investigate the relation between the cone $\mathcal{COP}^n$ of $n \times n$ copositive matrices and the approximating cone $\mathcal{K}_n^1$ introduced by Parrilo [Par00]. These cones are known to be equal for $n \leq 4$, and for $n \geq 5$ it is shown that they are not equal. Our result is based on the fact that $\mathcal{K}_n^1$ is not invariant under diagonal scaling while $\mathcal{COP}^n$ is. In particular it is shown that for any copositive matrix which is not the sum of a nonnegative and a positive semidefinite matrix, we can find a scaling which is not in $\mathcal{K}_n^1$. In fact, it can be shown that if all scaled versions of a matrix are contained in $\mathcal{K}_n^r$ for some fixed r, then the matrix must be in $\mathcal{K}_n^0$. For the specific 5×5 case, the more surprising result that we can scale any copositive matrix X into $\mathcal{K}_5^1$ is given. In particular it is shown that any scaling matrix D such that $(DXD)_{ii} \in \{0, 1\}$ for all i yields $DXD \in \mathcal{K}_5^1$. This provides a way to use the cone $\mathcal{K}_5^1$ to check if any order 5 matrix is copositive. Another consequence of this is a complete characterization of $\mathcal{COP}^5$ in terms of $\mathcal{K}_5^1$. Moreover, during Chapter 4 I provide an explicit way to construct a scaling matrix D for $A \in \mathcal{K}_n^1 \backslash \mathcal{K}_n^0$ such that $DAD \notin \mathcal{K}_n^1$, a result that was not included in [DDGH13b]. Another result given in Chapter 4 not included in this paper, is a negative result regarding the complexity of the problem of scaling arbitrary matrices in the opposite direction of the hierarchy of Parrilo cones. Finally I introduce the concept of non-decreasing scalings which also had not been discussed before in [DDGH13b]. At the end of Chapter 4 several conjectures are provided concerning scalings.

In *Chapter* 5 I suggest a copositive formulation for the graph isomorphism problem, that is, the problem of deciding whether or not two graphs are the same after a (possible) relabeling of the vertices. This problem is particular in the sense that its complexity is currently unknown. The hierarchies $\mathcal{C}_n^1$ and $\mathcal{K}_n^r$ are then applied to the copositive formulation and it is shown that an optimal solution always exists in one of these approximating cones for a finite r. This is particularly important for this problem because the graph isomorphism problem only has 'yes' and 'no' as answers. Finally, I rewrite the initial copositive formulation several times obtaining an LP formulation, as well as a possible method to construct a certificate for a polynomial time solution to the graph isomorphism problem.

Chapter 1

Introduction to copositive programming

In relatively recent years copositive programming has become a useful tool to deal with a wide variety of optimization problems, in particular with respect to the class of NP-hard problems, as the problem of checking for copositivity is itself NP-hard. This complexity result is due to Murty and Kabadi [MK87] who in fact showed that the problem of seeing whether a matrix is in the copositive cone is co-NP-complete. In Chapter 1 we will review some of the results obtained within copositive programming so far, providing a thorough introduction and overview of the field as well as giving a number of results that will be needed throughout this text. We shall start the introduction in Section 1.2 by focusing on some of the properties of the copositive cone and its dual, the completely positive cone. In Section 1.3 we will review a number of approximation techniques that are currently known for the copositive cone as well as the completely positive cone. In Section 1.4 we will give an overview of the field of copositive programming, describing optimization techniques as well as a number of problems known to be rewritable as a copositive (completely positive) program. However, first we will need to explain some of the notation used in this thesis, which we will do in Section 1.1. In particular, we will define most of the general notation that will be used throughout this text. More specific notation that will only be used once or twice, or only during a specific chapter or section, will further be defined throughout the rest of the text as it is needed. We included a nomenclature and an index at the end of the text to aid in the readability of this thesis.

1.1 Notation

We denote the space of n-dimensional real vectors by $\mathbb{R}^n$, rational vectors by $\mathbb{Q}^n$, vectors containing natural numbers by $\mathbb{N}^n$, and integer vectors by $\mathbb{Z}^n$. By $\mathbb{R}^n_+$ we denote the set of nonnegative real vectors, while $\mathbb{R}^n_{++}$ denotes the set of strictly positive real vectors. Moreover the space of $n \times m$ real matrices is denoted by $\mathbb{R}^{n\times m}$. Similar notation will be used for the other sets previously defined, e.g. $\mathbb{Q}^{n\times m}_+$ is the set of nonnegative rational $n\times m$ matrices. The set of symmetric real matrices of order n shall be denoted by $\mathbb{S}^n$ and the cone of $n\times n$ nonnegative real matrices by $\mathcal{N}^n$. We let Δ_n be the n-dimensional standard simplex. From now on we will omit the 'n' for any set if the dimension is equal to one, or when the dimension is clear from the context, in an effort to increase readability. All vectors will be printed in bold text to distinguish them from other variables and parameters. Likewise, we will generally use lowercase letters to denote scalars or vectors and capital letters to denote matrices. By $\mathbb{R}[\mathbf{x}]$ we denote the ring of polynomials in $\mathbf{x}$ with coefficients in $\mathbb{R}$.

The identity matrix and the all-ones matrix of order n will be written as I_n and E_n respectively. Similarly we let $\mathbf{e}_n$ be the vector of all-ones. Whenever we say $A \in \mathbb{R}^{n\times m}$ is nonnegative, or write $A \geq 0$, we mean that A is entry-wise nonnegative.

For any matrix $A \in \mathbb{R}^{n\times n}$, the operator $\mathrm{Tr} : \mathbb{R}^{n\times n} \to \mathbb{R}$ is the trace of A, i.e. the sum of diagonal elements of A. For the inner product we use the standard dot product, $\langle \mathbf{a}, \mathbf{b}\rangle = \mathbf{a}^\mathsf{T}\mathbf{b}$, when dealing with vectors $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$ and the trace inner product, $\langle A, B\rangle = \mathrm{Tr}(AB)$, when dealing with matrices $A, B \in \mathbb{S}^n$. The accompanying norms will be the Euclidean and the Frobenius norms respectively, that is $\|\mathbf{a}\| = \sqrt{\langle \mathbf{a}, \mathbf{a}\rangle}$ for $\mathbf{a} \in \mathbb{R}^n$ and $\|A\| = \sqrt{\langle A, A\rangle}$ for $A \in \mathbb{S}^n$. For $\mathbf{a} \in \mathbb{Z}^n_+$ we define $|\mathbf{a}| = \sum_{i=1}^n |a_i|$. For any set $K \subseteq \mathbb{R}^{n\times m}$ we denote its convex hull by $\mathrm{conv}(K)$ and its conic hull by $\mathrm{cone}(K)$. The closure of a set K shall be denoted $\mathrm{cl}(K)$ whereas its interior will be written as $\mathrm{int}(K)$. Then by $\dim(K)$ we denote the dimension of the set K.

The operator $\mathrm{Diag} : \mathbb{R}^n \to \mathbb{R}^{n\times n}$ turns a vector $\mathbf{d}$ into a diagonal matrix D where $D_{ii} = d_i, i = 1, \ldots, n$. Conversely $\mathrm{diag} : \mathbb{R}^{n\times n} \to \mathbb{R}^n$ returns a vector $\mathbf{a}$ such that $a_i = A_{ii}$, $i = 1, \ldots, n$, for any (not necessarily diagonal) matrix A. By a *scaling* of a matrix $A \in \mathbb{S}^n$ we mean the multiplication DAD, where $D := \mathrm{Diag}(d)$, $d \in \mathbb{R}_{++}$. Similarly, a matrix PAP, for some $A \in \mathbb{S}^n$ and $P \in \mathcal{P}_n$, is referred to as a *permutation* of A, where $\mathcal{P}_n$ is the set of all $n \times n$ permutation matrices. The operator $\mathrm{Vec} : \mathbb{R}^{n\times m} \to \mathbb{R}^{nm}$ turns a matrix $A = [a_1, \ldots, a_m]$, where a_i is the i^{th} column of A, into one long vector by stacking the columns of A, that is $\mathrm{Vec}(A) = [a_1^\mathsf{T}, \ldots, a_m^\mathsf{T}]^\mathsf{T}$. The Hadamard product of two matrices $A, B \in \mathbb{R}^{n\times m}$ is denoted as $A \circ B$. By $A \otimes B$ we mean the Kronecker product of the matrices A and B, for $A \in \mathbb{R}^{n\times m}$ and $B \in \mathbb{R}^{p\times q}$.

A graph is a pair $G = (V, E)$, where V (the vertex set) is some finite nonempty set and E (the edge set) is a set of 2-element subsets of V. By $\bar{G} = (V, \bar{E})$ we denote the complement graph of $G = (V, E)$ where $\{i, j\} \in \bar{E}$ if and only if $\{i, j\} \notin E$, for all $i, j \in V$.

When we refer to a *polynomial (respectively exponential) time* algorithm we mean that the maximum or worst-case computation time of the algorithm is polynomial (respectively exponential) in the encoding lengths of the inputs. Alternatively we say a problem is tractable if a polynomial time (or efficient) algorithm for it exists. By $\mathcal{O}(\bullet)$ we denote the usual big-O notation concerning limit behavior of functions, i.e. for any two functions f and

$g : \mathbb{R} \to \mathbb{R}$ we have $f(x) = \mathcal{O}(g(x))$ if there exists a constant $M \in \mathbb{R}_{++}$ such that $\limsup_{x\to\infty} |\frac{f(x)}{g(x)}| \le M$. For thorough introductions to complexity theory the reader is referred to Garey and Johnson [GJ79], and Grötschel, Lovász, and Schrijver [GLS88], and de Klerk [dK08]. During the remainder of this thesis we shall assume that the reader is familiar with basic complexity theory, including an understanding of encoding lengths.

1.2 The copositive cone and its dual

A symmetric matrix A is said to be *copositive* if the quadratic function $\mathbf{x}^\mathsf{T} A\mathbf{x}$ is nonnegative over the nonnegative orthant, that is for $\mathbf{x} \ge 0$. We will denote the cone of copositive matrices as follows.

Definition 1.1. The copositive cone is defined as

$$\mathcal{COP}^n = \{A \in \mathbb{S}^n \mid \mathbf{x}^\mathsf{T} A\mathbf{x} \ge 0 \text{ for all } \mathbf{x} \in \mathbb{R}^n_+\}. \tag{1.1}$$

The first ever mention of copositive forms, as well as the coinage of the term 'copositive' is credited to Motzkin [Mot52]. As mentioned before, the membership problem for the copositive cone, that is checking whether or not a given matrix is in $\mathcal{COP}^n$, is NP-hard. It turns out (see e.g. [BD08]) that in the definition of the copositive cone (1.1) we can limit $\mathbf{x}$ such that its Euclidean norm is equal to one, i.e.

$$\mathcal{COP}^n = \{A \in \mathbb{S}^n \mid \mathbf{x}^\mathsf{T} A\mathbf{x}\text{for all}\mathbf{x} \in \mathbb{R}^n_+,\ \|\mathbf{x}\| = 1\}.$$

We will use this fact later on in Chapter 2 to prove some complexity results. It should be noted that the definition of the copositive cone looks very similar to that of the well-studied *semidefinite cone*

$$\mathcal{S}^n_+ := \{A \in \mathbb{S}^n \mid \mathbf{x}^\mathsf{T} A\mathbf{x} \ge 0 \text{ for all } \mathbf{x} \in \mathbb{R}^n\}. \tag{1.2}$$

Yet copositive programming falls in the category of NP-hard problems whereas semidefinite programs are known to be solvable in polynomial time up to any desired accuracy. Note that this is currently the strongest known result with respect to the complexity of semidefinite programming and can be shown with the help of the so called *ellipsoid method* [YN76], [Sho77]. In particular it was the paper by Shor [Sho77] that gave the first explicit statement of the ellipsoid method as it is currently known. The ellipsoid method was also employed by Khachiyan [Kha79] to show that linear programming was in P. For a more thorough introduction to the ellipsoid method we refer the reader to Grötschel et al. [GLS88]. An important part of the ellipsoid method is the need for a so called *oracle* that decides whether or not a matrix is in (or within some guaranteed distance to) the feasible set. In Chapter 2 we will use a framework and machinery related to the ellipsoid method to prove, among other things, a complexity result for the dual of the copositive cone. For the semidefinite case an oracle is obtained with the help of the Cholesky factorization, which decomposes a matrix $A \in \mathcal{S}_+^n$ such that $A = BB^\mathsf{T}$, where $B \in \mathbb{R}^{n \times n}$ is a lower triangular matrix. The existence of such a decomposition, or in fact any decomposition $A = BB^\mathsf{T}$ with $B \in \mathbb{R}^{n \times k}$, provides an alternative definition of a positive semidefinite matrix, a fact easily seen from the definition of the semidefinite cone. In fact one of the convenient properties of the semidefinite cone is that it has several such equivalent definitions which we will list in the following proposition.

Proposition 1.2. *Let $A \in \mathbb{S}^n$, then the following statements are equivalent:*

i. the matrix A is positive semidefinite (resp. positive definite),

ii. all eigenvalues of A are nonnegative (positive),

iii. the determinant of every principal submatrix of A is nonnegative (positive),

iv. every principal submatrix of A is positive semidefinite (positive definite),

v. there exists a (nonsingular square) matrix B such that $A = BB^\mathsf{T}$.

For a proof of Proposition 1.2 we refer the reader to [BSM03, Theorem 1.10]. For the copositive cone, despite its similarity with the semidefinite cone in terms of its definition, very few conditions exist that are both necessary and sufficient. Moreover any such conditions, as we will see later on, are far less practical than their semidefinite counterparts. An exception to this observation is when $n \leq 4$, in which case we have that $\mathcal{COP}^n = \mathcal{S}_+^n + \mathcal{N}^n$. On the other hand, a couple of conditions similar to the conditions listed in Proposition 1.2 do exist, which instead are only necessary or only sufficient. We present these conditions concerning copositive matrices together with several others in the following proposition.

Proposition 1.3. *Let $A \in \mathcal{COP}^n$. Then we have:*

i. if P is a permutation matrix and D is a nonnegative diagonal matrix, then $PDADP^\mathsf{T} \in \mathcal{COP}^n$.

ii. every principal submatrix of A is also copositive.

iii. $(A)_{ii} \geq 0$ for all i.

iv. *if* $(A)_{ii} = 0$ *for some* i*, then* $(A)_{ij} \geq 0$ *for all* j*.*

v. $(A)_{ij} \geq -\sqrt{(A)_{ii}(A)_{jj}}$ *for all* i, j*.*

vi. *if there exists a strictly positive vector* $\mathbf{v}$ *such that* $\mathbf{v}^\mathsf{T} A\mathbf{v} = 0$*, then* A *is positive semidefinite,*

vii. *at least one eigenvalue, and moreover the sum of all eigenvalues, of* A *is nonnegative.*

For proofs of the properties (i) - (vi) we refer the reader to [Dia62] and [Dic13]. For a proof of (vii) see [HUS10], the latter paper moreover gives a condition in terms of so called Pareto eigenvalues that is both necessary and sufficient. It is a well known fact that the inequalities from (v) together with (iii) of Proposition 1.3 provide conditions for copositivity that are both necessary and sufficient for 2×2 matrices. In particular, this result follows immediately from the fact that $\mathcal{COP}^2 = \mathcal{S}^2_+ + \mathcal{N}^2$ combined with property (iii) of Proposition 1.2. Similar, allbeit more expensive, inequalities for checking copositivity can be obtained for the 3×3 [CS94] and 4×4 [PY93] case.

Note that although every principal submatrix of a copositive matrix is also copositive, the reverse does not hold. Sufficient, as well as necessary conditions in terms of principal submatrices, can however be constructed, as was shown by Gaddum [Gad58]. In this paper a recursive strategy for detecting copositivity is suggested based on a link to game theory, that is, the author obtains the following result.

Theorem 1.4. *Let* A *be a symmetric matrix and suppose that every principal submatrix of* A *is copositive. Then* A *is copositive if and only if*

$$\min_{\mathbf{x}\in\Delta_n} \max_{\mathbf{y}\in\Delta_n} \mathbf{y}^\mathsf{T} A\mathbf{x} = \max_{\mathbf{y}\in\Delta_n} \min_{\mathbf{x}\in\Delta_n} \mathbf{y}^\mathsf{T} A\mathbf{x} > 0. \tag{1.3}$$

The equality (1.3) can be verified via linear programming. Another recursive copositivity test requiring all principle submatrices to be copositive, is suggested by Cottle, Habetler, and Lemke [CHL70]. This result furthermore puts conditions on the so called adjugate matrix. For $A \in \mathbb{R}^{n\times n}$ the *adjugate matrix* of A is defined as an $n\times n$ matrix with $\mathrm{Adj}(A)_{ji} = (-1)^{i+j}\det(A(i,j))$, where $A(i,j)$ is a submatrix obtained from A by deleting row i and column j.

Theorem 1.5. *Let* A *be a symmetric matrix and suppose that every principle submatrix of* A *is copositive. Then* A *is copositive if and only if the determinant of* A *is nonnegative or the adjugate matrix of* A *contains a negative entry.*

Bomze [Bom87, Bom89] suggested yet another recursive way to check copositivity by introducing a criterion that looks similar to the Schur-complement for semidefinite matrices, albeit far more restricted.

Theorem 1.6. *Let* $\mathbf{b} \in \mathbb{R}^{n-1}$ *and* $C \in \mathbb{S}^{n-1}$. *Then* $\begin{bmatrix} a & \mathbf{b}^\mathsf{T} \\ \mathbf{b} & C \end{bmatrix}$ *is copositive if and only if*

i. $a \geq 0$ *and* C *is copositive, and*

ii. $\mathbf{x}^\mathsf{T}(aC - \mathbf{b}\mathbf{b}^\mathsf{T})\mathbf{x} \geq 0$ *for all* $\mathbf{x} \in \mathbb{R}^{n-1}_+$ *for which* $\mathbf{b}^\mathsf{T}\mathbf{x} \leq 0$.

An interesting remark regarding these conditions comes from [HUS10]. In that paper the authors observe that Theorem 1.6 boils down to checking copositivity of the 2×2 matrices

$$\begin{bmatrix} a & \mathbf{b}^\mathsf{T}\mathbf{x} \\ \mathbf{b}^\mathsf{T}\mathbf{x} & \mathbf{x}^\mathsf{T}C\mathbf{x} \end{bmatrix}$$

for all $\mathbf{x} \in \mathbb{R}^{n-1}_+$, which we can do by verifying a number of inequalities via properties (*iii*) and (*v*) of Proposition 1.3, as was mentioned before.

An algorithmic approach to check for copositivity was introduced in [BD08]. In this paper the authors introduce a formulation of the quadratic form $\mathbf{x}^\mathsf{T}A\mathbf{x}$ in barycentric coordinates with respect to the standard simplex. Let $\mathbf{v}_1, \ldots, \mathbf{v}_n$ be the vertices of a simplex. Then a point $\mathbf{x}$ in the simplex can be written as $\mathbf{x} = \sum_{i=1}^n \lambda_i \mathbf{v}_i$ with $\lambda_i \geq 0$ and $\sum_{i=1}^n \lambda_i = 1$. Then

$$\mathbf{x}^\mathsf{T}A\mathbf{x} = \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j \mathbf{v}_i^\mathsf{T} A \mathbf{v}_j \tag{1.4}$$

and so, for A to be copositive, a sufficient condition is

$$\mathbf{v}_i^\mathsf{T} A \mathbf{v}_j \geq 0 \text{ for all } i, j = 1, \ldots, n. \tag{1.5}$$

Partitioning the standard simplex into smaller simplices gives increasingly more and stronger constraints of this type. As it turns out, in the limit these type of constraints describe the copositive cone. A more general algorithm that also allows for a more general partitioning strategy is suggested in [BE12]. Moreover this paper provides a number of sufficient conditions for copositivity via linear and quadratic programming techniques.

The idea of using the standard simplex and barycentric coordinates has been considered before in [ACE95] where criteria for copositivity of matrices in $\mathbb{S}^n$ were constructed under the assumption that all principle submatrices (of order $n-1$) are copositive.

Another property that sets the copositive cone apart from the semidefinite cone concerns the dual cone.

Definition 1.7. Given a set $K \subseteq \mathbb{S}^n$, the *dual set* is defined by

$$K^* := \{A \in \mathbb{S}^n \mid \ \langle A, B\rangle \geq 0 \text{ for all } B \in K\}.$$

A set K is called *self-dual* if $K^* = K$.

Note that the semidefinite cone is a self-dual cone, a fact that can easily be seen from the definition given in (1.2) combined with the equivalent definition provided by (iv) of Proposition 1.2. To show the inclusion $(\mathcal{S}_+^n)^* \subseteq \mathcal{S}_+^n$, take $A \in (\mathcal{S}_+^n)^*$. We can take an arbitrary $\mathbf{b} \in \mathbb{R}^n$ and obtain

$$0 \leq \langle A, \mathbf{b}\mathbf{b}^\mathsf{T}\rangle = \mathbf{b}^\mathsf{T} A\mathbf{b}, \tag{1.6}$$

so $A \in \mathcal{S}_+^n$. To see the reverse inclusion, take $A \in \mathcal{S}_+^n$. To show that $A \in (\mathcal{S}_+^n)^*$, take $B \in \mathcal{S}_+^n$ and decompose as $B = \sum_{i=1}^k \mathbf{b}_i\mathbf{b}_i^\mathsf{T}$. Then

$$\langle A, B\rangle = \sum_{i=1}^k \langle A, \mathbf{b}_i\mathbf{b}_i^\mathsf{T}\rangle = \sum_{i=1}^k \mathbf{b}_i^\mathsf{T} A\mathbf{b}_i \geq 0, \tag{1.7}$$

which shows $A \in \mathcal{S}_+^n$.

Observe that the inequalities (1.6) and (1.7) in the proof given above stay valid when we let $A \in \mathcal{COP}^n$ as long as we furthermore demand that all vectors are nonnegative, i.e. $\mathbf{l}_1, \ldots, \mathbf{l}_k, \mathbf{b} \in \mathbb{R}_+^n$. This immediately gives us the definition of the dual of $\mathcal{COP}^n$ which is known as the *completely positive cone*. The completely positive cone was first mentioned roughly ten years after the introduction of the copositive cone in a paper by Hall Jr. [HJ62].

Definition 1.8. The completely positive cone is defined as

$$\mathcal{CP}^n = \{A \in \mathbb{S}^n \mid \ A = \sum_{i=1}^k \mathbf{b}_i\mathbf{b}_i^\mathsf{T}, \, \mathbf{b}_i \geq 0 \text{ for all } i\}$$

and is the dual of the copositive cone, that is $(\mathcal{CP}^n)^* = \mathcal{COP}^n$ and $\mathcal{CP}^n = (\mathcal{COP}^n)^*$. A decomposition $A = \sum_{i=1}^k \mathbf{b}_i\mathbf{b}_i^\mathsf{T}$ with $\mathbf{b}_i \in \mathbb{R}_+^n$ for all i is referred to as a *rank 1 decomposition*. More generally, for any matrix $A \in \mathcal{CP}^n$, $A = \sum_{i=1}^k B_i B_i^\mathsf{T}$ with $B_i \in \mathbb{R}^{n\times k}$ is referred to as a rank k decomposition of A.

As we will see in Chapter 2 the problem of deciding whether a matrix is completely positive or not is an NP-hard problem as well. Moreover, in Chapter 2 we will even give slightly stronger complexity results for both the copositive and the completely positive cone than the NP-hardness for their respective membership problems. Whether or not the membership problem for the completely positive cone is in the class NP as well is still an open question at the moment of writing. This question is connected to the well-studied problem of finding the so called CP-rank of a matrix.

Definition 1.9. Let $A \in \mathcal{CP}^n$, that is $A = \sum_{i=1}^{k} \mathbf{b}_i \mathbf{b}_i^{\mathsf{T}}$, where $\mathbf{b}_i \geq 0$ for $i = 1, \ldots, k$. Then the *CP-rank* of A is defined as the smallest possible such k.

For a detailed introduction to this topic the reader is referred to Chapter 3 of [BSM03]. It should be noted that bounds for the CP-rank exist that, moreover, are of polynomial nature with respect to n, the order of the matrices. This however does not solve the question whether or not membership for $\mathcal{CP}^n$ is in NP. The reason for this is that the entries of A and therefore also the entries of the vectors $\mathbf{b}_1, \ldots \mathbf{b}_k$ in Definition 1.9 are allowed to be in $\mathbb{R}_+$ and hence their encoding lengths need not be bounded with respect to n. An interesting development in this area is that the Drew-Johnson-Loewy conjecture that is mentioned in [BSM03] has since been almost entirely disproven, at least for the general case, in recent papers by Bomze et al. [BSU14a, BSU14b]. The Drew-Johnson-Loewy conjecture states that the upper bound to the CP-rank is equal to the lower bound $\lfloor \frac{n}{4} \rfloor$ that was found by Drew, Johnson, and Loewy [DJL94]. Bomze et al. [BSU14a] managed to find counterexamples on the boundary of the completely positive cone for $7 \leq n \leq 11$. They then managed to construct counterexamples for $n \geq 12$ as well in [BSU14b] obtaining a new bound for the CP-rank, $\frac{n^2}{2} + \mathcal{O}(n^{3/2})$. Shortly before these two papers, disproving the conjecture for the general case, the conjecture was shown to hold for the 5×5 case by Shaked-Monderer et al. [SMBJS13]. Whether or not the conjecture holds for 6×6 completely positive matrices is still an open question.

As with the copositive cone, equivalent definitions for complete positivity in the form of practical conditions that are both necessary and sufficient, do not currently seem to exist. Several necessary conditions have been found, some of which we will list in the following proposition.

Proposition 1.10. *Let $A, B \in \mathcal{CP}^n$. Then*

i if P is a permutation matrix and D is a nonnegative diagonal matrix, then $PDADP^{\mathsf{T}} \in \mathcal{CP}^n$, *ii every principal submatrix of A is completely positive,*

iii $A \otimes B$ is also completely positive.

For proofs we refer the reader to [BSM03]. As with the copositive cone, sufficient conditions exist as well, for example conditions based on diagonal dominance and comparison matrices. We say that a matrix A is *diagonally dominant (dd)* if $|(A)_{ii}| \geq \sum_{j \neq i} |(A)_{ij}|$ for all i. Furthermore, a matrix A is called *scaled diagonally dominant (sdd)* if there exists a positive diagonal matrix D such that DAD is diagonally dominant. For dd and sdd matrices we have the following theorem.

Theorem 1.11. *Let $A \in \mathbb{S}^n$ be a scaled diagonally dominant matrix with nonnegative diagonal entries. Then $A \in \mathcal{S}_+^n$.*

Proof. If A is sdd then DAD is dd for some diagonal matrix D with positive diagonal entries. The result now follows from [BSM03, Proposition 1.8] in combination with (v) of Proposition 1.2. $\square$

Finally, the *comparison matrix* of A is defined as

$$M(A)_{ij} = \begin{cases} |(A)_{ij}| & \text{if } i = j, \\ -|(A)_{ij}| & \text{if } i \neq j. \end{cases}$$

Now there is the following results on complete positivity.

Proposition 1.12. *Let $A \in \mathcal{N}^n$. Then*

i if A is (scaled) diagonally dominant then A is completely positive,

ii if the comparison matrix $M(A)$ is positive semidefinite, then A is completely positive.

iii if A is positive semidefinite with rank equal to r, and if A has an $r \times r$ principle submatrix that is diagonal, then A is completely positive.

For proofs of (i) and (ii) we again refer the reader to [BSM03]. It should be noted that (i) is only shown for diagonally dominant matrices in [BSM03]. The fact that this property also holds for scaled diagonally dominant matrices can easily be seen from the fact that if $DAD \in \mathcal{CP}^n$ for a diagonal matrix D with positive diagonal then $A \in \mathcal{CP}^n$. Moreover, several more technical and more specific sufficient conditions can be found in that same book, including several conditions based on graph theoretical concepts. Condition (iii) is a result by Shaked-Monderer [SM09]. This result has slightly been improved upon by

Kalofolias and Gallopoulos [KG12] who manage to construct an explicit rank 2 decomposition for this case.

Dong, Lin and Chu [DLC14] developed a heuristic method to check complete positivity. Their method works as long as the CP-rank of the matrix being evaluated is equal to the rank of that matrix. Another approach is suggested by Dickinson and Dür [DD12], who present an algorithm that deals with sparse matrices. They prove that this algorithm can be used to determine complete positivity, as well as give an explicit factorization, for tridiagonal and acyclic matrices in linear time.

Both the copositive and the completely positive cone are so called *proper cones*, i.e. both cones are full dimensional, closed, convex, and pointed. For proofs we refer the reader to [Dic13, Chapter 5]. The interior of the copositive cone is described, analogous to the interior of the positive semidefinite cone, as

$$\operatorname{int}(\mathcal{COP}^n) = \{A \in \mathbb{S}^n \mid \mathbf{x}^\mathsf{T} A\mathbf{x} > 0 \text{ for all } \mathbf{x} \in \mathbb{R}_+ \setminus \{0\}\}. \tag{1.8}$$

The interior of the completely positive cone, however, is not quite as straight forward. It was investigated in [DS08] where the authors obtained the following description:

$$\operatorname{int}(\mathcal{CP}^n) = \{A \in \mathbb{S}^n \mid A = BB^\mathsf{T},\ B = [B_1|B_2] \text{ where } B_1 > 0 \text{ nonsingular, } B_2 \geq 0\}$$

Several alternative definitions of the interior of the completely positive cone are given in [Dic13, Chapter 7].

A full description of the copositive cone in terms of its extreme rays is currently unknown. For $n \leq 4$, however, we know that $\mathcal{COP}^n = \mathcal{S}_+^n + \mathcal{N}^n$, for which all extreme rays are known from [HN63]. Attempts to go further in this direction have been made by Baumert [Bau65, Bau66, Bau67], Baston [Bas69], and Ycart [Yca82]. The paper by Hall and Newman [HN63] gives a full description of the extreme rays of the completely positive cone. If we let $\operatorname{Ext}(K)$ be the set of extreme rays for some set $K \in \mathbb{R}^{n\times m}$, then Hall and Newman show that

$$\operatorname{Ext}(\mathcal{CP}^n) = \{\mathbf{a}\mathbf{a}^\mathsf{T} \mid \mathbf{a} \in \mathbb{R}_+^n \setminus \{\mathbf{0}\}\}. \tag{1.9}$$

Moreover it was shown in [Dic11] that all such extreme rays are also exposed rays. For the case of 5×5 matrices a full description of the extreme rays of the copositive cone is given by Hildebrand [Hil12]. Details of this description will be given in Chapter 4 of this thesis. For $n \geq 6$ a full characterization of the extreme rays of $\mathcal{COP}^n$ is not known. However a number of extreme rays of the copositive cone have been found and investigated by [HN63, Dic11]. In particular [Dic13] gives an overview of the currently known results in this area and also manages to give a full description of the maximal faces of the copositive cone. For further surveys on properties of the copositive and completely positive cone, the reader is referred to Hiriart-Urruty and Seeger [HUS10] and Berman and Shaked-Monderer [BSM03].

1.3 Approximation hierarchies

Over the last several years a lot of effort has been spent to obtain approximations for the copositive cone, as well as for the completely positive cone. The necessity for such approximations seems two-fold. On the one hand the complexity results regarding these cones imply that no efficient oracles exist unless P= NP. In fact, as we will see later in Chapter 2 even the problem of deciding whether a matrix is within a guaranteed distance of the copositive or completely positive cone is NP-hard. In practice this means that none of the general conic optimization frameworks can be used to solve copositive programs for even medium sized instances. On the other hand no non-efficient easy to use state-of-the-art solvers for copositive programs are currently available either. Hence, in order to deal with copositive programs, especially for medium to large instances, we are left to approximations which are generally more tractable and moreover, as we shall see below, can make use of existing frameworks and solvers.

Some obvious approximations include the cones $\mathcal{S}^n_+$, $\mathcal{S}^n_+ + \mathcal{N}^n$ and $\mathcal{S}^n_+ \cap \mathcal{N}^n$. The latter is referred to in the literature as the *doubly nonnegative cone*. From the definitions we get the following simple inclusions

$$\mathcal{CP}^n \subseteq \mathcal{S}^n_+ \cap \mathcal{N}^n \subseteq \mathcal{S}^n_+ + \mathcal{N}^n \subseteq \mathcal{COP}^n,$$

where we get equality for $\mathcal{CP}^n \subseteq \mathcal{S}^n_+ \cap \mathcal{N}^n$ and $\mathbb{S}^n + \mathcal{N}^n \subseteq \mathcal{COP}^n$ if and only if $n \leq 4$. Examples showing that this is not true anymore for $n \geq 5$, are the so called *Horn-matrix* [HN63],

$$H = \begin{pmatrix} 1 & -1 & 1 & 1 & -1 \\ -1 & 1 & -1 & 1 & 1 \\ 1 & -1 & 1 & -1 & 1 \\ 1 & 1 & -1 & 1 & -1 \\ -1 & 1 & 1 & -1 & 1 \end{pmatrix} \in \mathcal{COP}^5 \setminus (\mathcal{S}^5_+ + \mathcal{N}^5), \tag{1.10}$$

and the matrix

$$H^* = \begin{pmatrix} 1 & 1 & 0 & 0 & 1 \\ 1 & 2 & 1 & 0 & 0 \\ 0 & 1 & 2 & 1 & 0 \\ 0 & 0 & 1 & 2 & 1 \\ 1 & 0 & 0 & 1 & 6 \end{pmatrix} \in (\mathcal{S}_+^5 \cap \mathcal{N}^5) \setminus \mathcal{CP}^5, \tag{1.11}$$

which is taken from [BSM03]. For a more thorough description of the difference between the cones $\mathcal{CP}^n$ and $\mathcal{S}_+^n \cap \mathcal{N}^n$, especially for the 5×5 case, we refer the reader to [BAD09] and [DA13]. Moreover [DA13] provides a construction that separates matrices in $(\mathcal{S}_+^5 \cap \mathcal{N}^5) \setminus \mathcal{CP}^5$ from $\mathcal{CP}^n$, as well as such matrices of arbitrary size having some block structure.

1.3.1 Approximations via simplicial partitioning

Note that the cones mentioned above can be handled using the well studied framework of semidefinite programming. In fact many of the existing approximations for copositive programs involve the construction of cones that make use of semidefinite programming techniques. An exception to this is the technique introduced in [BD08], which is based on the conditions in (1.5). The set described by these conditions gets refined in the approach described by Bundfuss and Dür resulting in a hierarchy of cones approximating $\mathcal{COP}^n$ from within, a result made explicit in [BD09]. Assume we partition the standard simplex $r \in \mathbb{Z}_+$ times into a number of smaller simplices, then denote by $\mathbb{P}_r \subseteq \Delta_n$ and $\mathbb{E}_r$ the sets of vertices and edges of these simplices respectively. The authors then introduce the hierarchy of inner approximations of $\mathcal{COP}^n$

$$\mathcal{I}_r^n := \left\{ A \in \mathbb{S}^n \;\middle|\; \begin{array}{l} \mathbf{v}^\mathsf{T} A \mathbf{v} \geq 0 \text{ for all } \mathbf{v} \in \mathbb{P}_r \\ \mathbf{u}^\mathsf{T} A \mathbf{v} \geq 0 \text{ for all } (\mathbf{u}, \mathbf{v}) \in \mathbb{E}_r \end{array} \right\}, \quad r \in \mathbb{Z}_+.$$

They furthermore show that these cones converge to the copositive cone, i.e. $\operatorname{cl}\left(\bigcup_{r\in\mathbb{Z}_+} \mathcal{I}_r^n\right) = \mathcal{COP}^n$. The authors suggest outer approximations for the copositive cone as well by relaxing the set of constraints (1.5) to $\mathbf{v}^\mathsf{T} A \mathbf{v} \geq 0$ instead. That is they introduce the hierarchy of outer approximations of $\mathcal{COP}^n$

$$\mathcal{O}_r^n := \left\{ A \in \mathbb{S}^n \mid \mathbf{v}^\mathsf{T} A \mathbf{v} \geq 0 \text{ for all } \mathbf{v} \in \mathbb{P}_r \right\}, \quad r \in \mathbb{Z}_+.$$

These cones are shown to converge to $\mathcal{COP}^n$ as well which means that $\bigcap_{r\in\mathbb{Z}_+} \mathcal{O}_r^n = \mathcal{COP}^n$. These approximations for the copositive cone can in turn be used to construct outer and inner approximations for the completely positive cone through duality. Explicit formulation of these cones approximating $\mathcal{CP}^n$ are provided in [BD09].

In general, nontrivial outer approximations for the copositive cone, and hence inner approximations of the completely positive cone, seem hard to find. They do exist, however their quality varies. Observe that we can trivially create outer approximations of the copositive cone, similar to the one described above, by simply defining a cone

$$\{A \in \mathbb{S}^n \mid \mathbf{x}^\mathsf{T} A\mathbf{x} \geq 0 \text{ for all } \mathbf{x} \in U\}$$

where $U \subset \mathbb{R}^n_+$. For example we can define U as the set of binary vectors.

The idea of using simplicial partitioning is not unique to copostive programming though, and has been applied to other types of nonlinear optimization problems, see e.g. [Hor76, Kea78, TH88, Tuy88, Hor97]. For a more detailed description of such techniques the reader is referred to [Dic13, Chapter 9] and [HPT95].

1.3.2 The polyhedral approximation cones $\mathcal{C}_n^r$

The construction of so called approximation hierarchies like the ones we saw in the previous section, is an approach that is often employed to approximate the copositive or completely positive cone. A number of such hierarchies have been proposed in the literature, however unlike the hierarchies introduced in [BD09], most of them hinge on two basic concepts. First the condition $\mathbf{x}^\mathsf{T} A\mathbf{x} \geq 0$ for all $\mathbf{x} \in \mathbb{R}^n_+$ is replaced by some sufficient condition $\mathbf{x}^\mathsf{T} A\mathbf{x} \in \phi$ for nonnegativity that is closed under multiplication. Doing so automatically creates a cone

$$\Phi_n := \{A \in \mathbb{S}^n \mid \mathbf{x}^\mathsf{T} A\mathbf{x} \in \phi\} \subseteq \mathcal{COP}^n$$

that is a (not necessarily strict) subset of the copositive cone. This cone is then expanded upon by multiplying the quadratic term a number of times with some function f that is positive over $\mathbb{R}^n_+ \setminus \{\mathbf{0}\}$ for which the condition $f \in \phi$ trivially holds, often chosen as $f(\mathbf{x}) = (\sum_{i=1}^n x_i)$ in existing hierarchies for $\mathcal{COP}^n$. This produces a hierarchy of cones

$$\Phi_n^r := \{A \in \mathbb{S}^n \mid f(\mathbf{x})^r \mathbf{x}^\mathsf{T} A\mathbf{x} \in \phi\} \subseteq \mathcal{COP}^n, \quad r \in \mathbb{Z}_+.$$

Note that $\Phi_n^r \subseteq \Phi_n^{r+1}$ for all $r \in \mathbb{Z}_+$ due to the fact that $f \in \phi$ by assumption while ϕ is closed under multiplication. Furthermore note that for any positive function f over $\mathbb{R}^n_+ \setminus \{\mathbf{0}\}$ it holds that $f(\mathbf{x}) \cdot \mathbf{x}^\mathsf{T} A\mathbf{x} \geq 0$ for all $\mathbf{x} \in \mathbb{R}^n_+$ implies $\mathbf{x}^\mathsf{T} A\mathbf{x} \geq 0$ for all $\mathbf{x} \in \mathbb{R}^n_+$, ensuring that $\Phi_n^r \subseteq \mathcal{COP}^n$ for all $r \in \mathbb{Z}_+$.

An obvious sufficient condition for any polynomial over $\mathbb{R}^n_+$ to be nonnegative, is to demand that all its coefficients are nonnegative, which can be verified using linear programming techniques. This idea was employed by de Klerk and Pasechnick in [dKP02] to construct a hierarchy of polyhedral cones $\mathcal{C}_n^0, \mathcal{C}_n^1, \dots$ that converge to $\mathcal{COP}^n$ which they defined as follows:

$$\mathcal{C}_n^r := \Big\{ A \in \mathbb{S}^n \mid \Big(\sum_{i=1}^n x_i\Big)^r \mathbf{x}^\mathsf{T} A\mathbf{x} \text{ has nonnegative coefficients } \Big\}. \tag{1.12}$$

Note that if $\big(\sum_{i=1}^n x_i\big)^r \mathbf{x}^\mathsf{T} A\mathbf{x}$ has nonnegative coefficients for some $r \geq 1$ then $\big(\sum_{i=1}^n x_i\big)^{r+1} \mathbf{x}^\mathsf{T} A\mathbf{x} = \big(\sum_{i=1}^n x_i\big)\big(\sum_{i=1}^n x_i\big)^r \mathbf{x}^\mathsf{T} A\mathbf{x}$ must have nonnegative coefficients as well so that $\mathcal{C}_n^r \subseteq \mathcal{C}_n^{r+1} \subseteq \mathcal{COP}^n$. The convergence of the cones $\mathcal{C}_n^r$ that was claimed above is a direct result from a theorem by Pólya.

Theorem 1.13 (Pólya [Pól28])**.** *Let f be a homogeneous polynomial which is positive on the simplex $\Delta = \{\mathbf{z} \in \mathbb{R}^n_+ \mid \sum_{i=1}^n z_i = 1\}$. Then for sufficiently large N, all the coefficients of the polynomial $(\sum_{i=1}^n z_i)^N f(\mathbf{z})$ are nonnegative.*

This result implies that $\bigcup_{r \in \mathbb{Z}_+} \mathcal{C}_n^r = \operatorname{int}(\mathcal{COP}^n)$ so that $\operatorname{cl}(\bigcup_{r \in \mathbb{Z}_+} \mathcal{C}_n^r) = \mathcal{COP}^n$.

1.3.3 The Parrilo r-cones

The concept of creating hierarchies of cones approximating $\mathcal{COP}^n$ as described at the start of section 1.3.2, was first introduced by Parrilo in his thesis [Par00]. The sufficient condition he used was that of sum of squares, which can be dealt with via the framework of semidefinite programming. A polynomial $f \in \mathbb{R}[\mathbf{x}]$ is said to be *sum of squares (SOS)* if there exist a finite number of polynomials $g_1, \dots, g_k \in \mathbb{R}[\mathbf{x}]$ such that $f = \sum_{i=1}^k g_i^2$. The connection between SOS and semidefinite programming is made explicit by the following theorem.

Theorem 1.14 (See e.g. [Lau08])**.** *Let $f \in \mathbb{R}[\mathbf{x}]$ be a polynomial of degree $2d$. Furthermore let $(\mathbf{x}^{[\leq d]})$ be the vector containing all monomials of degree at most d. Then*

$$f \text{ is SOS} \iff \text{there exists a } B \in \mathcal{S}_+^k \text{ such that } f = (\mathbf{x}^{[\leq d]})^\mathsf{T} B(\mathbf{x}^{[\leq d]}). \tag{1.13}$$

Note that for general polynomials of degree $2d$, the matrix B as in Theorem 1.14 is of order k, where

$$k = \sum_{i=1}^{d} \binom{n+i-1}{i}.$$

In order to apply Theorem 1.14 we require a polynomial to have an even degree. Hence, for $r \in \mathbb{Z}_+$ and taking into account the discussion above the Parrilo r-cone as introduced in [Par00] is defined as

$$\mathcal{K}_n^r := \Big\{ A \in \mathbb{S}^n \mid \Big(\sum_{i=1}^{n} x_i^2 \Big)^r (\mathbf{x} \circ \mathbf{x})^\mathsf{T} A (\mathbf{x} \circ \mathbf{x}) \text{ is SOS } \Big\}. \tag{1.14}$$

The squares of the variables $x_1, \ldots, x_n$ that appear in (1.14) can be seen as a replacement of the non-negativity of $\mathbf{x}$. More practically it means the resulting polynomial $(\sum_{i=1}^n x_i^2)^r (\mathbf{x} \circ \mathbf{x})^\mathsf{T} A (\mathbf{x} \circ \mathbf{x})$ has an even degree so that we can use Theorem 1.14 to verify membership of $\mathcal{K}_n^r$ using semidefinite programming.

As with the cones $\mathcal{C}_n^r$ it can immediately be seen that $\mathcal{K}_n^r \subseteq \mathcal{K}_n^{r+1} \subseteq \mathcal{COP}^n$, for every $r \in \mathbb{Z}_+$. Next, observe that any polynomial over $\mathbb{R}_+^n$ with nonnegative coefficients can trivially be written as a sum of squares so that $\mathcal{C}_n^r \subseteq \mathcal{K}_n^r$ which furthermore implies that $\mathrm{cl}(\bigcup_{r \in \mathbb{Z}_+} \mathcal{K}_n^r) = \mathcal{COP}^n$.

As it turns out $\mathcal{K}_n^0 = \mathcal{S}_+^n + \mathcal{N}^n$. Moreover recall that for $n \leq 4$ we have $\mathcal{COP}^n = \mathcal{S}_+^n + \mathcal{N}^n$, so that in fact $\mathcal{K}_n^0 = \mathcal{COP}^n$ for $n \leq 4$. Clearly this also means that $\mathcal{K}_n^0 \neq \mathcal{COP}^n$ for $n \geq 5$ due to the example of the Horn matrix (1.10) mentioned before. In fact as we shall see later in Chapter 4 there exist no other cases where $\mathcal{K}_n^r = \mathcal{COP}^n$ other than for when $r = 0$ and $n \leq 4$. Note that determining whether or not a polynomial with degree $2d$ is sum of squares requires solving a rather large positive semidefinite program of order $\sum_{i=1}^d \binom{n+d-1}{n-1}$ together with a large number of linear equality constraints. Even in the case of checking membership for $\mathcal{K}_n^r$, where we have a homogeneous polynomial of degree $2r+4$, the resulting positive semidefinite matrix B is still of the order $\binom{n+r+1}{n-1}$. This means that for increasing values of r and n, the cones $\mathcal{K}_n^r$ become computationally expensive very quickly.

1.3.4 The SOS approximation cones $\mathcal{Q}_n^r$

Another approximation hierarchy closely related to the Parrilo r-cones, is the series of cones introduced by Peña et al. [PVZ07]. They again use sum of squares as a sufficient condition for non-negativity, as in the Parrilo cones, but rather in a more restrictive way. For $\mathbf{x} \in \mathbb{R}_+^n$ and $\boldsymbol{\beta} \in \mathbb{Z}_+^n$ denote $\mathbf{x}^{\boldsymbol{\beta}} = \prod_{i=1}^n x_i^{\beta_i}$

and recall that $|\boldsymbol{\beta}| = \sum_{i=1}^{n} \beta_i$, then for any $r \in \mathbb{Z}_+$ and $n \in \mathbb{N}$ Peña et al. define their hierarchy of cones as

$$\mathcal{Q}_n^r = \Big\{ A \in \mathbb{S}^n \mid \Big(\sum_{i=1}^{n} x_i \Big)^r \mathbf{x}^\mathsf{T} A \mathbf{x} = \sum_{\boldsymbol{\beta} \in \mathbb{Z}_+^n, |\boldsymbol{\beta}| = r} \mathbf{x}^{\boldsymbol{\beta}} \mathbf{x}^\mathsf{T} B_\beta \mathbf{x},\ B_{\boldsymbol{\beta}} \in \mathcal{S}_+^n + \mathcal{N}^n \Big\}. \tag{1.15}$$

As with the Parrilo cones we again have the inclusions $\mathcal{Q}_n^r \subseteq \mathcal{Q}_n^{r+1} \subseteq \mathcal{COP}^n$ as well as the property that $\mathrm{cl}(\bigcup_{r \in \mathbb{Z}_+} \mathcal{K}_n^r) = \mathcal{COP}^n$, due to the fact that obviously $\mathcal{C}_n^r \subseteq \mathcal{Q}_n^r$ for all $r \in \mathbb{Z}_+$. Again, recall that $\mathcal{S}_+^n + \mathcal{N}^n = \mathcal{K}_n^0$. Then immediately from Theorem 1.14 it follows that $\mathbf{x}^\mathsf{T} B_\beta \mathbf{x}$, and hence $\mathbf{x}^\beta \mathbf{x}^\mathsf{T} B_\beta \mathbf{x}$, is sum of squares for every $\boldsymbol{\beta} \in \mathbb{Z}_+^n$, $|\boldsymbol{\beta}| = r$. In other words, the cones $\mathcal{Q}_n^r$ break the polynomial $\big(\sum_{i=1}^n x_i\big)^r \mathbf{x}^\mathsf{T} A \mathbf{x}$ into a number of smaller bits which they then demand to be sum of squares representable. This way the authors restrict the more general SOS condition from the Parrilo cones $\mathcal{K}_n^r$ to a number of smaller SOS conditions. Furthemore note that the terms $\mathbf{x}^{\boldsymbol{\beta}}$ in the summation on the right-hand side of the equality in (1.15), are simply the monomials of degree r produced by $(\sum_{i=1}^n x_i)^r$. From this it can be seen that we can equivalently express membership of the cones $\mathcal{Q}_n^r$ via a set of *linear matrix inequalities (LMIs)*. In particular, it was noted in [PVZ07] that in order to decide whether $A \in \mathcal{Q}_n^r$, we can instead check whether there exist matrices $B_{\boldsymbol{\beta}} \in \mathcal{S}_+^n$, $\boldsymbol{\beta} \in \mathbb{Z}_+^n$ where $|\boldsymbol{\beta}| = r$ such that

$$\Big(\sum_{i=1}^{n} x_i \Big)^r \mathbf{x}^\mathsf{T} A \mathbf{x} \trianglerighteq \sum_{\boldsymbol{\beta} \in \mathbb{Z}_+^n, |\boldsymbol{\beta}| = r} \mathbf{x}^{\beta} \mathbf{x}^\mathsf{T} B_{\boldsymbol{\beta}} \mathbf{x}, \tag{1.16}$$

where "$\trianglerighteq$" indicates that the coefficients on the left-hand side are greater than or equal to those on the right-hand side for any given monomial. For example, to verify membership of $\mathcal{Q}_n^2$ for some matrix $A \in \mathbb{S}^n$ we obtain the LMIs

$$\Big(\sum_{i=1}^{n} x_i \Big)^2 \mathbf{x}^\mathsf{T} A \mathbf{x} \trianglerighteq \sum_{1 \le i \le j \le n} x_i x_j \mathbf{x}^\mathsf{T} B_{ij} \mathbf{x}, \quad B_{ij} \in \mathcal{S}_+^n. \tag{1.17}$$

Furthermore it is noted in [PVZ07] that we can use Proposition 9 from [PVZ06] to equivalently write the Parrilo cones as

$$\mathcal{K}_n^r = \Big\{ A \in \mathbb{S}^n \mid \Big(\sum_{i=1}^{n} x_i \Big)^r \mathbf{x}^\mathsf{T} A \mathbf{x} = \sum_{\boldsymbol{\beta} \in \mathbb{Z}_+^n, |\boldsymbol{\beta}| \le r+2} \mathbf{x}^{\boldsymbol{\beta}} \mathbf{x}^\mathsf{T} B_{\boldsymbol{\beta}} \mathbf{x},\ B_{\boldsymbol{\beta}} \in \mathcal{S}_+^n + \mathcal{N}^n \Big\}. \tag{1.18}$$

From this we immediately see that $\mathcal{Q}_n^r \subseteq \mathcal{K}_n^r$ for every $r \in \mathbb{Z}_+$ and $n \in \mathbb{N}$. Moreover, from (1.18) it becomes clear that the number of variables and LMIs for the $\mathcal{K}_n^r$ cones grows at a much faster rate, with both n and r, than for the cones $\mathcal{Q}_n^r$. Comparing to the example given above where $r = 2$, i.e. (1.17), we get a whole extra matrix variable B of size $\binom{n+1}{2} \times \binom{n+1}{2}$ for $\mathcal{K}_n^2$, i.e. denoting by $(\mathbf{x}^{[d]})$ the vector containing all monomials of degree exactly d we obtain the LMIs

$$\Big(\sum_{i=1}^{n} x_i\Big)^2 \mathbf{x}^\mathsf{T} A\mathbf{x} \trianglerighteq \sum_{1\le i\le j\le n} x_i x_j \mathbf{x}^\mathsf{T} B_{ij}\mathbf{x} + (\mathbf{x}^{[2]})^\mathsf{T} B(\mathbf{x}^{[2]}).$$

1.3.5 Approximating cones via DSOS and SDSOS

As we have seen so far, particularly in Theorem 1.14, a polynomial can be determined to be sum of squares using semidefinite programming. In theory this means that for a given value of r we can determine membership for both $\mathcal{K}_n^r$ and $\mathcal{Q}_n^r$ for any matrix (up to any desired accuracy) in polynomial time with respect to the size n of this matrix. In practice however, especially for large scale optimization, semidefinite programming is not nearly as efficient as linear and second order cone programming. Moreover we can see from (1.15) and (1.18) that the size of the semidefinite programs increases quite rapidly with both n and r for these sum of squares problems. As a consequence of these two issues alternatives are needed to deal with large scale optimization. Several approximations have been suggested for this purpose based on linear or second order cone programming, an example are the cones $\mathcal{C}_n^r$ described above. Other examples include the hierarchies of cones introduced by Ahmadi and Majumdar [AM14]. Their approximations are based on relaxations of the SOS constraints. Ahmadi and Majumdar introduce two concepts called diagonally dominant sum of squares and scaled diagonaly dominant sum of squares. A polynomial $f \in \mathbb{R}[\mathbf{x}]$ is said to be *diagonally dominant sum of squares (DSOS)* if there exist monomials $m_i \in \mathbb{R}[\mathbf{x}]$ and nonnegative constants α_i and $\beta_{i,j}$, $i, j = 1, \ldots, k$ for some $k \in \mathbb{N}$, such that

$$f(\mathbf{x}) = \sum_{i=1}^{k} \alpha_i m_i^2 + \sum_{i=1}^{k}\sum_{i=1}^{k} \beta_{i,j}(m_i \pm m_j)^2.$$

Similarly, a polynomial $f \in \mathbb{R}[\mathbf{x}]$ is said to be *scaled diagonaly dominant sum of squares (SDSOS)* if there exist monomials $m_i \in \mathbb{R}[\mathbf{x}]$ and constants $\alpha_i \in \mathbb{R}_+$ and $\beta_i, \gamma_i \in \mathbb{R}$, $i = 1, \ldots, k$ for some $k \in \mathbb{N}$ such that

$$f(\mathbf{x}) = \sum_{i=1}^{k} \alpha_i m_i^2 + \sum_{i=1}^{k} \sum_{i=1}^{k} (\beta_i m_i \pm \gamma_j m_j)^2.$$

Analogue to the approximation hierarchies presented above the authors then introduce a hierarchy of sets *r-DSOS* and *r-SDSOS*. That is,

$$\text{r-DSOS} := \Big\{ f \in \mathbb{R}[\mathbf{x}] \mid \big(\sum_{i=1}^{n} x_i^2\big)^r f \text{ is DSOS } \Big\},$$

and

$$\text{r-SDSOS} := \Big\{ f \in \mathbb{R}[\mathbf{x}] \mid \big(\sum_{i=1}^{n} x_i^2\big)^r f \text{ is SDSOS } \Big\}.$$

Restricting the polynomials to those of the form $f = \mathbf{x}^\mathsf{T} A \mathbf{x}$ for $A \in \mathbb{S}^n$, combined with Theorem 2.1 of [AM14] the authors propose the following hierarchies of cones approximating $\mathcal{COP}^n$:

$$\mathrm{DD}_n^r := \Big\{ A \in \mathbb{S}^n \mid \big(\sum_{i=1}^{n} x_i^2\big)^r (\mathbf{x} \circ \mathbf{x})^\mathsf{T} A (\mathbf{x} \circ \mathbf{x}) = (\mathbf{x}^{[r+1]})^\mathsf{T} B (\mathbf{x}^{[r+1]}),\ B \text{ is dd} \Big\}, \tag{1.19}$$

and

$$\mathrm{SDD}_n^r := \Big\{ A \in \mathbb{S}^n \mid \big(\sum_{i=1}^{n} x_i^2\big)^r (\mathbf{x} \circ \mathbf{x})^\mathsf{T} A (\mathbf{x} \circ \mathbf{x}) = (\mathbf{x}^{[r+1]})^\mathsf{T} B (\mathbf{x}^{[r+1]}),\ B \text{ is sdd} \Big\}, \tag{1.20}$$

where we recall that $(x^{[r+1]})$ is the vector containing all monomials of degree $r+1$. Note the relationship of (1.19) and (1.20) with Theorem 1.11 and Theorem 1.14, making the concept that these cones are a relaxation of the sum of squares conditions, as presented above, explicit. Moreover from this observation we immediately get that

$$\mathrm{DD}_n^r \subseteq \mathrm{SDD}_n^r \subseteq \mathcal{K}_n^r$$

From [AM14, Theorem 2.2] we know that DD_n^r is polyhedral while SDD_n^r has a second order cone representation. Comparing the cones (1.19) and (1.20) to the alternative formulation (1.18) of $\mathcal{K}_n^r$, we furthermore see that apart from having an easier matrix constraint (i.e. diagonally dominant and scaled diagonally dominant versus positive semidefinite) the number of matrix variables is significantly smaller as well. Then, as with the other hierarchies it can easily be seen that $\mathrm{DD}_n^r \subseteq \mathrm{DD}_n^{r+1}$ and $\mathrm{SDD}_n^r \subseteq \mathrm{SDD}_n^{r+1}$ for all $r \in \mathbb{Z}_+$.

1.3.6 Other approximations for $\mathcal{COP}^n$ and $\mathcal{CP}^n$

A generalization that contains the hierarchies $\mathcal{C}_n^r$, $\mathcal{Q}_n^r$, and $\mathcal{K}_n^r$ has been suggested by Dobre and Vera [DV13]. They define the hierarchy of cones

$$\mathbb{H}_{k,n}^r := \left\{ A \in \mathbb{S}^n \mid \Big(\sum_{i=1}^n x_i\Big)^r \mathbf{x}^\mathsf{T} A \mathbf{x} = \sum_{j=r+2-k}^{r+2} \sum_{\boldsymbol{\beta}\in\mathbb{N}^n, |\boldsymbol{\beta}|=j} \mathbf{x}^{\boldsymbol{\beta}} \sigma_{\boldsymbol{\beta}}(\mathbf{x}),\ \sigma_{\boldsymbol{\beta}}(\mathbf{x}) \in \Sigma_{r+2-|\boldsymbol{\beta}|} \right\}$$

where for any $d \in \mathbb{N}$, Σ_d denotes the set of sum of squares polynomials of degree d. Given $n \in \mathbb{N}$, for any nondecreasing sequence k_r, $r = 0, 1, 2, \dots$ they obtain

$$\mathbb{H}_{k_r,n}^r \subseteq \mathbb{H}_{k_{r+1},n}^{r+1} \subseteq \mathcal{COP}^n$$

as well as $\mathrm{cl}(\bigcup_{r\in\mathbb{Z}_+} \mathbb{H}_{k_r,n}^r) = \mathcal{COP}^n$. Moreover they get that $\mathcal{C}_n^r = \mathbb{H}_{0,n}^r$ and $\mathcal{Q}_n^r = \mathbb{H}_{2,n}^r$ while $\mathcal{K}_n^r = \mathbb{H}_{k,n}^r$ for any $k \geq r+2$.

Another method to construct approximations for the copositive cone strongly related to the concept of sum of squares is that based on the theory of moments. Although most of this approximation theory is set up in a more general framework of possibly restricted nonnegative polynomials over some set K, it can be tailored to the copositive cone. That is, we can reduce these techniques to the copositive setting by restricting the polynomials to have degree at most two and by taking K as the nonnegative orthant. One of the most well known results in this area is the so called Lasserre hierarchy which was introduced in its general form in [Las11]. Approximation hierarchies, specifically for the copositive and completely positive cone, were then presented in [Las14]. These result were later extended by Dickinson [Dic13, Corollary 11.3]. In fact, this extension explicitly shows the link between the theory of moments and sum of squares. Another way this link becomes apparent is via the dual of the cones $\mathcal{K}_n^r$, see e.g. [Dic13, Section 11.6]. As the theory of moments is outside of the scope of this text we refer the interested reader to [GL07] and [Dic13, Chapter 10] for an introduction to this topic.

Finally, two more hierarchies of outer approximations of the completely positive cone were suggested by Dong [Don13] and are constructed using symmetric tensors. A *tensor* of order r and dimension n is an object for which each entry is sepcified using r indices each of which takes values in $\{1, \ldots, n\}$. Let $\sigma : \{1, \ldots, r\} \to \{1, \ldots, r\}$ be a permutation. Then T is a *symmetric tensor* of order r if $T(\sigma(i_1), \ldots, \sigma(i_r)) = T(i_1, \ldots, i_r)$ for all permutations σ. Observe that a symmetric tensor for $r = 2$ is simply a symmetric matrix. Denoting the set of symmetric tensors of order r and dimension n by $\mathbb{S}_n^r$, Dong furthermore defines the operators $\mathbf{Slices}(\bullet)$ and $\mathbf{Collapse}(\bullet)$. In particular $\mathbf{Slices}(Z)$ is the set of $n \times n$ symmetric matrices that can be produced from $Z \in \mathbb{S}_n^{r+2}$ by fixing exactly r indices of Z. Then $\mathbf{Collapse}(Z)$ is the sum of all such matrices produced by $\mathbf{Slices}(Z)$. The approximation hierarchies are then obtained by first rewriting the completely positive cone as

$$\mathcal{CP}^n = \{A \in \mathbb{S}^n \mid \exists Z \in \mathbb{S}_n^{r+2}, \mathbf{Slices}(Z) \subseteq \mathcal{CP}^n, A = \mathbf{Collapse}(Z)\}.$$

The condition $\mathbf{Slices}(Z) \subseteq \mathcal{CP}^n$ in the equation above is then relaxed to the weaker conditions $\mathbf{Slices}(Z) \subseteq \mathcal{N}^n$ and $\mathbf{Slices}(Z) \subseteq \mathcal{S}_+^n \cap \mathcal{N}^n$ respectively. Doing so produces the approximation hierarchies

$$\mathcal{T}_n^r = \{A \in \mathbb{S}^n \mid \exists Z \in \mathbb{S}_n^{r+2}, \mathbf{Slices}(Z) \subseteq \mathcal{N}^n, A = \mathbf{Collapse}(Z)\}, \tag{1.21}$$

and

$$\mathcal{TD}_n^r = \{A \in \mathbb{S}^n \mid \exists Z \in \mathbb{S}_n^{r+2}, \mathbf{Slices}(Z) \subseteq \mathcal{S}_+^n \cap \mathcal{N}^n, A = \mathbf{Collapse}(Z)\}. \tag{1.22}$$

Furthermore, it is shown in [Don13] that the cones $\mathcal{T}_n^r$ and $\mathcal{C}_n^r$ are mutually dual, as are the cones $\mathcal{TD}_n^r$ and $\mathcal{Q}_n^r$. A result following directly from these duality statements combined with Theorem 1.13 is that the cones $\mathcal{T}_n^r$ and $\mathcal{TD}_n^r$ converge to $\mathcal{CP}^n$, i.e. $\bigcap_{r \in \mathbb{Z}_+} \mathcal{T}_n^r = \mathcal{CP}^n$ as well as $\bigcap_{r \in \mathbb{Z}_+} \mathcal{TD}_n^r = \mathcal{CP}^n$.
Finally, in light of all these approximation hierarchies we introduce the lifting rank of a matrix $A \in \mathbb{S}^n$ that we will particularly make use of in Chapter 3 in an effort to shorten and simplify notation.

Definition 1.15. Let K be some cone in $\mathbb{R}^{n \times n}$. For any matrix $A \in K$ and any hierarchy of sets $(\mathbb{Y}_n^r)_{r \in \mathbb{Z}_+}$ approximating K we define the *lifting rank* of A as,

$$r^*_{\mathbb{Y}_n^r}(A) = \min\{r \in \mathbb{Z}_+ \mid A \in \mathbb{Y}_n^r\}. \tag{1.23}$$

When there does not exist an r for which $A \in \mathbb{Y}_n^r$ we define $r^*_{\mathbb{Y}_n^r}(A) = \infty$.

As an example, for any matrix $A \in \mathcal{COP}^4$ we have $r^*_{\mathcal{K}_4^r}(A) = 0$.

1.4 Copositive (Completely Positive) Programming

When we speak of a copositive, or alternatively a completely positive program, we are referring to a linear program over either the copositive or completely positive cone respectively. That is, a *copositive program* is of the form

$$\begin{aligned} &\min \langle C, X\rangle \\ &\text{s.t.}\ \langle A_i, X\rangle = b_i \quad i = 1, \dots, m \\ &\qquad X \in \mathcal{COP}^n, \end{aligned} \tag{1.24}$$

where the dual problem, a *completely positive program*, is

$$\begin{aligned} &\max \mathbf{b}^{\mathsf{T}}\mathbf{y} \\ &\text{s.t.}\ C - \sum_{i=1}^{m} y_i A_i \in \mathcal{CP}^n,\ \mathbf{y} \in \mathbb{R}^n, \end{aligned} \tag{1.25}$$

for given matrices A_i, $C \in \mathbb{R}^{n\times n}$ and vector $\mathbf{b} \in \mathbb{R}^n$. Note that both of these programs are convex due to the fact that both $\mathcal{COP}^n$ and $\mathcal{CP}^n$ are convex cones. It was then shown by Eichfelder and Jahn [EJ08] that if Slater's condition is satisfied, the Karush-Kuhn-Tucker optimality conditions hold, and moreover we get strong duality as long as both problems are feasible and at least one of them is strictly feasible, i.e. allows for a feasible solution in the interior of either the copositive or completely positive cone. At the moment of writing no exact algorithms exist to solve copositive or completely positive programs to optimality. We can however solve strictly feasible copositive programs up to any desired accuracy using the algorithm introduced by Bundfuss and Dür [BD09]. This algorithm extends the idea presented in [BD08] that we described earlier in this chapter, see Section 1.3.1. In particular the authors propose an adaptive approach where moreover the partitioning of the simplex no longer happens uniformly, but is instead guided in an adaptive manner via the objective function. A heuristic approach that approximates completely positive problems from within via a descent method was introduced by Bomze et al. [BJR11].

Another way of dealing with copositive programs is via approximations, for example via any of the hierarchies that were mentioned in Section 1.3. In particular, substituting $\mathcal{COP}^n$ in (1.24) by any subset of $\mathcal{COP}^n$ one immediately obtains upper bounds to this minimization problem. The dual of any such subset of $\mathcal{COP}^n$ furthermore provides us with an outer approximation to the completely positive cone, allowing us to obtain upper bounds for the completely positive maximization problem (1.25) as well. Similarly, lower bounds for (1.24) and upper bounds for (1.25) can be obtained via outer approximations of $\mathcal{COP}^n$ and inner approximations of $\mathcal{CP}^n$ respectively.

As mentioned before, the motivation for copositive programming comes from the fact that a number of NP-hard problems can be written as copositive or completely positive programs. Combined with the approximation techniques described in Section 1.3, a number of new bounds and other interesting results have been obtained in recent years. The remainder of this section will be dedicated to reviewing several such results.

One of the most straightforward examples comes from a paper by Bomze et al. [BDdK+00], where the authors rewrite the standard quadratic problem as a completely positive program. The standard quadratic problem is defined as

$$\min\{\mathbf{x}^{\mathsf{T}} Q \mathbf{x} \mid \mathbf{e}^{\mathsf{T}}\mathbf{x} = 1,\, \mathbf{x} \in \mathbb{R}^n_+\}. \tag{1.26}$$

By writing $\mathbf{x}^{\mathsf{T}} Q \mathbf{x} = \langle Q, \mathbf{x}\mathbf{x}^{\mathsf{T}} \rangle$ one immediately sees from the definition of $\mathcal{CP}^n$ that $\mathbf{x}\mathbf{x}^{\mathsf{T}}$ is a completely positive rank 1 matrix. Hence one can define the following completely positive relaxation of (1.26)

$$\min\{\langle Q, X\rangle \mid \langle X, E\rangle = 1,\, X \in \mathcal{CP}^n\}. \tag{1.27}$$

Note that (1.27) is a linear program over the cone $\mathcal{CP}^n$. In particular, the linearity of the objective function implies that the optimal solution to (1.27) is attained at an extreme point of $\mathcal{CP}^n$. As mentioned before, see (1.9), the extreme rays of the completely positive cone are exactly its rank 1 matrices, and so the completely positive formulation (1.27) is in fact exact. Bomze and de Klerk [BdK02] later used this completely positive formulation, together with the approximation hierarchies $\mathcal{C}_n^r$ from (1.12) and $\mathcal{K}_n^r$ from (1.14), to construct a polynomial time approximation scheme for the standard quadratic problem.

One of the most well known results in copositive programming is Burer's result concerning binary quadratic optimization from [Bur09]. In this paper it was shown that the binary quadratic program

$$\begin{aligned} \min\ & \mathbf{x}^{\mathsf{T}} Q \mathbf{x} + \mathbf{c}^{\mathsf{T}}\mathbf{x} \\ \text{s.t.}\ & \mathbf{a}_i^{\mathsf{T}}\mathbf{x} = b_i \qquad (i = 1, \dots, m) \\ & \mathbf{x} \geq 0 \\ & x_j \in \{0, 1\} \qquad (j \in B) \end{aligned} \tag{1.28}$$

can be written equivalently as

$$
\begin{aligned}
\min\ & \langle Q, X\rangle + \mathbf{c}^\mathsf{T}\mathbf{x} \\
\text{s.t. }\ & \mathbf{a}_i^\mathsf{T}\mathbf{x} = b_i && (i = 1,\dots,m) \\
& \mathbf{a}_i^\mathsf{T} X \mathbf{a}_i = b_i^2 && (i = 1,\dots,m) \\
& x_j = X_{jj} && (j \in B) \\
& \begin{bmatrix} 1 & \mathbf{x}^\mathsf{T} \\ \mathbf{x} & X \end{bmatrix} \in \mathcal{CP}^{n+1}
\end{aligned} \tag{1.29}
$$

under the assumption that the conditions $\mathbf{x} \geq 0$ and $\mathbf{a}_i^\mathsf{T}\mathbf{x} = b_i$, $i = 1,\dots,m$, imply that $x_j \leq 1$ for all $j \in B$. This condition is referred to by Burer as *the key condition* and can, without loss of generality, always be achieved by introducing slack variables s_i and adding the constraints $x_i + s_i = 1$, $i = 1,\dots,n$ to the completely positive program (1.29). Moreover, Burer showed that under the additional assumption that there exists a $\mathbf{y} \in \mathbb{R}^m$ such that $\mathbf{h} := \sum_{i=1}^m y_i a_i \geq 0$ and $\sum_{i=1}^m y_i b_i = 1$, one can eliminate the variables $\mathbf{x}$ from (1.29) altogether. These extra assumptions, in other words, reduce (1.29) from an optimization problem over $\mathcal{CP}^{n+1}$ to an optimization problem over $\mathcal{CP}^n$, that is (1.29) is equivalent to

$$
\begin{aligned}
\min\ & \langle Q, X\rangle + \mathbf{c}^\mathsf{T} X\mathbf{h} \\
\text{s.t. }\ & \mathbf{a}_i^\mathsf{T} X\mathbf{h} = b_i && (i = 1,\dots,m) \\
& \mathbf{a}_i^\mathsf{T} X \mathbf{a}_i = b_i^2 && (i = 1,\dots,m) \\
& [X\mathbf{h}]_j = X_{jj} && (j \in B) \\
& X \in \mathcal{CP}^n \\
& \mathbf{h}^\mathsf{T} X\mathbf{h} = 1.
\end{aligned}
$$

Another interesting note on the feasible set of the completely positive program (1.29) was published by Bomze and Jarre [BJ10]. In particular they provide a direct way to prove equivalence of (1.28) and (1.29) which is significantly shorter than that in [Bur09]. The linear case, that is that of mixed binary linear programs, was treated by Natarajan et al [NTZ11] in a robust optimization setting. For this type of robust optimization problem the authors managed to obtain a completely positive formulation, and moreover provide an implicit connection between their formulation and the theory of moments.

Another well known result for copositive programming is that regarding the stability number. Given a graph $G = (V, E)$, the *stability number* is the maximum cardinality of a subset of vertices such that no two of them are connected by an edge. A copositive formulation for this problem, containing only one variable and one matrix constraint, was proposed by de Klerk and

Pasechnick [dKP02]. Let α_G denote the stability number of a graph $G = (V, E)$, and let A_G be its adjacency matrix, then the authors show that

$$\alpha_G = \min_{\lambda \in \mathbb{R}} \{\lambda \mid \lambda(I + A_G) - E \in \mathcal{COP}^n\}. \tag{1.30}$$

Replacing $\mathcal{COP}^n$ with $\mathcal{C}_n^r$ they then obtain a hierarchy of polyhedral approximations for the stability number, which they denote by $\zeta^{(r)}(G)$, $r \in \mathbb{Z}_+$. The authors show that $\lfloor \zeta^{(\alpha_G^2)} \rfloor = \alpha_G$. Moreover, it was conjectured by De Klerk and Pasechnik that this polyhedral approximation hierarchy would be able to find the stability number for $r = \alpha_G$, i.e. $r^*_{\mathcal{C}_n^r}(\alpha_G(I + A_G) - E) = \alpha_G$. This conjecture was shown to be true by Laurent and Gvozdenović [GL07] for graphs with stability number at most 8. The dual of the copositive formulation (1.30) is given by

$$\alpha_G = \max\{\langle E, X\rangle \mid \langle A_G + I, X\rangle = 1,\ X \in \mathcal{CP}^n\}. \tag{1.31}$$

By replacing the completely positive constraint $X \in \mathcal{CP}^n$ in (1.31) with the doubly nonnegative constraint $X \in \mathcal{S}_+^n \cap \mathcal{N}^n$, we obtain the famous upper bound known as the Lovász-Schrijver ϑ' number. A copositive formulation for the stability number of infinite graphs was suggested by Dobre et al. [DDV14].

Closely related to the stability number is the so called chromatic number of a graph. The *chromatic number*, χ_G, for a graph G is defined as the smallest number of colors needed to color the vertices of G such that any two vertices with the same color are not connected by an edge. The relation with the stability number comes from the fact that in any coloring of a graph each color defines a stable set. A copositive formulation for this problem was found by Gvozdenović and Laurent [GL08]. For a graph G, let G_k be the graph obtained by taking the Cartesian product of the complete graph on k vertices and G. Then the following is a copositive formulation of the chromatic number:

$$\chi_G = \max_{y,z \in \mathbb{R}} \{y \mid \tfrac{1}{n^2}(k - y)E_{nk} + z(n(I_{nk} + A_{G_k}) - E_{nk}) \in \mathcal{COP}^{nk},\ k = 1, \dots, n\}.$$

Improvements on the Lovász ϑ number in the direction of the chromatic number were obtained by Dukanovic and Rendl [DR10]. Given a graph $G = (V, E)$ with $|E| = m$, the authors define for $\mathbf{y} \in \mathbb{R}^{|E|}$ the operator $A_G^\mathsf{T}\mathbf{y} := \sum_{\{i,j\} \in E} y_{ij}(\mathbf{e}_i\mathbf{e}_j^\mathsf{T} + \mathbf{e}_j\mathbf{e}_i^\mathsf{T})$, where $\mathbf{e}_i$ denotes the i^{th} unit vector. Then they present the following completely positive program

$$\min_{z \in \mathbb{R}_+} \{z \mid zI_m + A_G^\mathsf{T}\mathbf{y} - E_m \in \mathcal{S}_+^m,\ zI_m + A_G^\mathsf{T}\mathbf{y} \in \mathcal{CP}^m\}.$$

The authors furthermore show that the optimal value of this completely positive formulation is equal to the so called fractional chromatic number. The *fractional chromatic number* is similar to the chromatic number, except that vertices are assigned sets of colors rather than a single color. Moreover the authors use the Parrilo cones $\mathcal{K}_n^r$ from (1.14) to obtain approximations, which they then simplify for vertex transitive graphs. They also show that the Lovász ϑ number can be improved upon significantly for specific Hamming graphs.

The quadratic assignment problem was reformulated as a copositive program by Povh and Rendl [PR09]. Let $A, B \in \mathbb{S}^n$ and $C \in \mathbb{R}^{n \times n}$. Then the quadratic assignment problem is defined as

$$\min \Big\{ \langle X, AXB + C \rangle \mid X \in \mathcal{P}_n \Big\}. \tag{1.32}$$

Next, Povh and Rendl define $c := \mathrm{Vec}(C)$ and show that (1.32) can be formulated as the following copositive program

$$\max_{S,T \in \mathbb{S}^n, v \in \mathbb{R}} \Big\{ \mathrm{Tr}(S) + \mathrm{Tr}(T) + n^2 v \mid A \otimes B + \mathrm{Diag}(c) - I \otimes S - T \otimes I - vE_{n^2} \in \mathcal{COP}^{n^2} \Big\}.$$

This is done by adding redundant constraints, and with the help of Lagrangian duality. A more detailed description of this method is given in Chapter 5, where we will use this technique to find a copositive formulation for the graph isomorphism problem.

Another problem studied by Povh and Rendl [PR07] is the so called graph partitioning problem. The *graph partitioning problem* is the problem of partitioning the vertices of an edge weighted graph, $G_w = (V, E, W)$, into sets V_1,V_2,V_3 of prescribed cardinalities $m_1, m_2, m_3 \in \mathbb{Z}_+$ respectively, such that the total weight of all edges between V_1 and V_2 is minimized. For their completely positive formulation of this problem the authors define matrices $E_n^{ij} := \mathbf{e}_i \mathbf{e}_j^\mathsf{T}$. Furthermore, let A_{G_w} be the weighted adjacency matrix of the weighted graph G_w. Then for any such graph G_w the graph partitioning problem can equivalently be written as

$$
\begin{aligned}
\min\ & \tfrac{1}{2}\Big\langle E_3^{12} + E_3^{21} \otimes A_{G_w}, Y \Big\rangle \\
\text{s.t.}\ & \tfrac{1}{2}\Big\langle \big(E_3^{ij} + E_3^{ji}\big) \otimes I_n, Y \Big\rangle = m_i \delta_{ij} && 1 \le i \le j \le 3 \\
& \Big\langle E_3 \otimes E_n^{ii} \Big\rangle = 1 && 1 \le i \le n \\
& \Big\langle \Big(\sum_{j=1}^{3} E_3^{ij}\Big) \otimes \Big(\sum_{i=1}^{n} E_n^{ij}\Big), Y \Big\rangle = m_i && 1 \le i \le 3,\ 1 \le j \le n \\
& \tfrac{1}{2}\Big\langle \big(E_3^{ij} + E_3^{ji}\big) \otimes E_n, Y \Big\rangle = m_i m_j && 1 \le i \le j \le n \\
& Y \in \mathcal{CP}^{3n}.
\end{aligned}
\tag{1.33}
$$

The completely positive formulation (1.33) is then relaxed to obtain a semidefinite program, whose optimal value turns out to be equal to the spectral bound that was found by Helmberg et al. [HRMP95]. Contrary to that bound however, the semidefinite formulation obtained by Povh and Rendl allows for a natural way to add extra constraints without losing tractability of the problem. The authors use this observation to obtain a tightened version of the semidefinite program and hence the spectral bound for the graph partitioning problem. For small instances the authors show that this tightened version can in fact give significant improvements over the spectral bound.

For other surveys on the field of copositive programming we refer the reader to surveys by Dür (2010) [Dür10], Burer (2012) [Bur12], Bomze (2012) [Bom12], Dickinson (2013) [Dic13], and Bomze, Schachinger and Uchida (2012) [BSU12].

Chapter 2

Complexity of membership for the completely positive cone and its dual[1]

[1]Published as [DG14] Peter J.C. Dickinson and Luuk Gijben. On the computational complexity of membership problems for the completely positive cone and its dual. *Computational Optimization and Applications*, 57(2):403-415,2014.

As noted in the introduction, it has been proven by Murty and Kabadi [MK87] that the strong membership problem for the copositive cone is co-NP-complete. As a result, no polynomial time algorithms will exist for solving general copositive programs unless P = NP. Motivated by this result of Murty and Kabadi we consider the question of whether the strong membership problem for the completely positive cone is NP-hard. It is commonly believed that this is true and in fact this is to such an extent that many researchers in the field believe it to have already been proven. However the technical details concerning this question have never been considered in order to supply a proof for this result and confirm that the suspicion concerning the complexity is correct. It was conjectured by Jarre and Schmallowsky [JS09] that the strong membership problem for the completely positive cone is an NP-complete problem. In this chapter we provide a proof that the membership problem for the completely positive cone is indeed NP-hard and it is left as an open question as to whether this problem in NP as well. Note in particular the discussion concerning the CP-rank in the introduction of this thesis regarding this question. Moreover we will give a proof for the result that even weak membership for the completely positive cone is NP-hard. In particular, we prove this result in Section 2.3 by establishing a polynomial time Turing reduction from the stable set problem, a known NP-complete problem (see for example Garey and Johnson [GJ79, Section 3.1.3]) to the weak membership problem for the completely positive cone. Introductions to the stable set problem are provided by Bondy and Murty [BM76, Chapter 7] and Schrijver [Sch03, Chapter 64]. We recall from Grötschel, Lovász, and Schrijver [GLS88] that for two problems Π_1, Π_2, a *(polynomial time) Turing Reduction* from Π_1 to Π_2 is an algorithm A_1 which solves Π_1 using a hypothetical subroutine A_2 which solves Π_2 such that if A_2 is a polynomial time algorithm then so is A_1. A special case of Turing reductions is the many-one reduction where the algorithm A_2 is only called once. In these reductions it is vital that the encoding length of the input to the algorithm A_2 is polynomial in the input to the algorithm A_1, and providing such polynomial inputs will be the main work during this chapter.

In this chapter we will also provide an alternative proof for NP-hardness of the strong membership problem for the copositive cone in Section 2.2 together with a proof that even the weak membership problem for the copositive cone is an NP-hard problem.

We will however start off this chapter by providing some definitions and known results, as well as a technical lemma, in Section 2.1.

2.1 Problems for convex sets

We start off this section by recalling several definitions from Grötschel, Lovász, and Schrijver [GLS88], which are extended for the space of symmetric matrices. In this section we let $\mathbb{X}$ be equal to either $\mathbb{R}^n$ or $\mathbb{S}^n$ and correspondingly let $\mathcal{Q}$ be equal to either $\mathbb{Q}^n$ or $(\mathbb{Q}^{n\times n} \cap \mathbb{S}^n)$ respectively.

Definition 2.1. Let $K \subseteq \mathbb{X}$ and let $\varepsilon > 0$. Then we define,

$$\begin{aligned} \mathrm{S}\,(K,\varepsilon) &:= \{x \in \mathbb{X} \mid \|x - y\| \leq \varepsilon \text{ for some } y \in K\}, \\ \mathrm{S}\,(K,-\varepsilon) &:= \{x \in \mathbb{X} \mid \ \mathrm{S}\,(\{x\},\varepsilon) \subseteq K\}. \end{aligned}$$

When $K = \{a\}$ we shall write $\mathrm{S}\,(a,\varepsilon) := \mathrm{S}\,(\{a\},\varepsilon)$.

Note that we have the following relation, $\mathrm{S}\,(K,-\varepsilon) \subseteq K \subseteq \mathrm{S}\,(K,\varepsilon)$. Hence $\mathrm{S}\,(K,-\varepsilon)$ and $\mathrm{S}\,(K,\varepsilon)$ can be seen as inner and outer approximations of K respectively.

We now consider the following problems for a set $K \subseteq \mathbb{X}$.

Definition 2.2. The Strong Membership Problem (MEM). Let $K \subseteq \mathbb{X}$. Given an instance $y \in \mathcal{Q}$, decide that either

1. $y \in K$, or
2. $y \notin K$.

Definition 2.3. The Weak Membership Problem (WMEM). Let $K \subseteq \mathbb{X}$. Given an instance $(y,\delta) \in \mathcal{Q} \times \mathbb{Q}_{++}$, decide that either

1. $y \in \mathrm{S}\,(K,\delta)$, or
2. $y \notin \mathrm{S}\,(K,-\delta)$.

Note that for $K \neq \emptyset, \mathbb{X}$ and $\delta > 0$, we have that $\mathrm{S}\,(K,\delta) \setminus \mathrm{S}\,(K,-\delta) \neq \emptyset$. Therefore, in general, for some instances either answer would be valid.

Definition 2.4. The Weak Validity Problem (WVAL). Let $K \subseteq \mathbb{X}$. Given an instance $(c,\gamma,\varepsilon) \in \mathcal{Q} \times \mathbb{Q} \times \mathbb{Q}_{++}$, either

1. decide that $\langle c, x\rangle \leq \gamma + \varepsilon$ for all $x \in \mathrm{S}\,(K,-\varepsilon)$, or

2. decide that $\exists y \in \mathrm{S}(K,\varepsilon)$ for which $\langle c,y\rangle \geq \gamma-\varepsilon$.

Note that again, in general, for some instances either answer would be valid.

We now consider what WVAL tells us about the values of $\langle c,x\rangle$ for $x\in K$, rather than just for x in some approximation of K. However, before we do this we first define a special type of convex body.

Definition 2.5. Consider a convex set $K\subseteq \mathbb{X}$ with the following properties,

1. $N=\dim \mathbb{X}$,
2. $\exists R\in\mathbb{Q}_{++}$ such that $K\subseteq \mathrm{S}(0,R)$, and
3. $\exists r\in\mathbb{Q}_{++}$, $a_0\in\mathcal{Q}$ such that $\mathrm{S}(a_0,r)\subseteq K$.

Then K is called an a_0-centered convex body which is denoted as the quintuple $(K;N,R,r,a_0)$.

Lemma 2.6. *Consider WVAL with K being a convex body $(K;N,R,r,a_0)$ as defined in Definition 2.5. If we assume that $\varepsilon<r$ then we have the following,*

1. *$\langle c,x\rangle\leq\gamma+\varepsilon$ for all $x\in\mathrm{S}(K,-\varepsilon)$ implies that*
$$\langle c,z\rangle \leq \left(\gamma+\varepsilon-\frac{\varepsilon}{r}\langle c,a_0\rangle\right)\Big/\left(1-\frac{\varepsilon}{r}\right) \qquad \textit{for all } z\in K.$$
2. *$\exists y\in\mathrm{S}(K,\varepsilon)$ for which $\langle c,y\rangle\geq\gamma-\varepsilon$ implies that*
$$\exists z\in K \textit{ such that } \langle c,z\rangle \geq \gamma-(1+\|c\|)\varepsilon.$$

Proof. We shall prove both points of the theorem separately.

1. Let $z_0\in K$, then
$$\mathrm{conv}\left(\{z_0\}\cup\mathrm{S}(a_0,r)\right)\subseteq K,$$
Therefore, if we let $z_\theta=(1-\theta)z_0+\theta a_0$, then we have that $\mathrm{S}(z_\theta,\theta r)\subseteq K$ for all $0\leq\theta\leq 1$. Hence in particular $z_{\varepsilon/r}\in\mathrm{S}(K,-\varepsilon)$. We now get that
$$\begin{aligned}\langle c,z_0\rangle &= \left(\langle c,z_\theta\rangle-\theta\langle c,a_0\rangle\right)/\left(1-\theta\right) \qquad \text{for all } 0\leq\theta<1\\ &= \left(\langle c,z_{\varepsilon/r}\rangle-\frac{\varepsilon}{r}\langle c,a_0\rangle\right)\Big/\left(1-\frac{\varepsilon}{r}\right)\\ &\leq \left(\gamma+\varepsilon-\frac{\varepsilon}{r}\langle c,a_0\rangle\right)\Big/\left(1-\frac{\varepsilon}{r}\right).\end{aligned}$$

2. Let $y \in \mathrm{S}(K,\varepsilon)$ such that $\langle c, y\rangle \geq \gamma - \varepsilon$ and let $z \in \mathrm{S}(y,\varepsilon) \cap K$ $(\neq \emptyset)$, then

$$\begin{aligned}\langle c, z\rangle &\geq \min\{\langle c, u\rangle \mid u \in \mathrm{S}(y,\varepsilon)\}\\ &= \min\{\langle c, y\rangle + \varepsilon\langle c, v\rangle \mid v \in \mathrm{S}(0,1)\}\\ &= \langle c, y\rangle - \|c\|\varepsilon\\ &\geq \gamma - (1+\|c\|)\varepsilon.\end{aligned}$$ □

We will now consider how the problems in this section are related to each other. It is immediately apparent that an oracle for MEM would provide us with an oracle for WMEM. For $\mathbb{X} = \mathbb{R}^n$, we now connect WMEM and WVAL using the following lemma by Yudin and Nemirovski [YN76].

Lemma 2.7. *For $\mathbb{X} = \mathbb{R}^n$, there exists an algorithm that solves WVAL for every quintuple $(K; n, R, r, a_0)$ given by a weak membership oracle. This algorithm is oracle-polynomial time with respect to the encoding lengths of the quintuple and the instance for WVAL.*

It is relatively easy to show that this theorem also holds for $\mathbb{X} = \mathbb{S}^n$. However, when doing this, care must be taken to maintain rationality. One way to do this would be to define the one-to-one mapping "svec" which maps $\mathbb{S}^n$ to $\mathbb{R}^N$, where $N = \dim(\mathbb{S}^n) = \frac{1}{2}n(n+1)$, by stacking the elements of the upper triangle, i.e.

$$\operatorname{svec}(X) = (x_{11},\ x_{12},\ x_{22},\ x_{13}, \ldots,\ x_{1k},\ x_{2k}, \ldots,\ x_{kk}, \ldots,\ x_{nn})^{\mathsf{T}}.$$

For any two matrices $C, X \in \mathbb{S}^n$, we then get that

$$\begin{aligned}&\langle C, X\rangle = \langle \operatorname{svec}((2E - I)\circ C), \operatorname{svec}(X)\rangle,\\ &\tfrac{2}{3}\|X\| \leq \|\operatorname{svec}(X)\| \leq \|X\|.\end{aligned}$$

The important things about this mapping are that it is linear, rationality is maintained and the norms are linearly related. (An alternative definition for the svec mapping which often appears in the literature has the off-diagonal elements multiplied by $\sqrt{2}$. Although this has the advantage of maintaining the inner product and norms, we would not maintain rationality, as required.) Using this, along with Lemma 2.7, it is now trivial to prove the following theorem.

Theorem 2.8. *For $\mathbb{X}$ equal to $\mathbb{R}^n$ or $\mathbb{S}^n$, there exists an algorithm that solves WVAL for every quintuple $(K; N, R, r, a_0)$ given by a weak membership oracle. This algorithm is oracle-polynomial time with respect to the encoding lengths of the quintuple and the instance for WVAL.*

2.2 The Copositive Cone

If a graph G contains a stable set of size λ then it is easy to see that it contains a stable set of size $t \in \mathbb{Z}_{++}$ for all $t \leq \lambda$. Recall that we denote by $\alpha_G \in \mathbb{Z}_{++}$ the stability number of a graph G. Note that G contains a stable set of size $t \in \mathbb{Z}_{++}$ if and only if $\alpha_G \geq t$. Combining this observation with (1.30) and (1.31) we get the following alternative formulations of the stability number of a graph G

$$\begin{aligned} \alpha_G &= \min \left\{ \lambda \mid \big((I + A_G)\lambda - E \big) \in \mathcal{COP} \right\}, && (2.1) \\ &= \max \left\{ \langle E, X \rangle \mid \langle I + A_G, X \rangle = 1,\ X \in \mathcal{CP} \right\}, && (2.2) \\ &= \max \left\{ \langle E, X \rangle \mid \langle I + A_G, X \rangle \leq 1,\ X \in \mathcal{CP} \right\}, && (2.3) \end{aligned}$$

It can now be seen that the stable set problem is Turing reducible to MEM for the copositive cone.

Lemma 2.9. *The graph G contains a stable set of size $t \in \mathbb{Z}_{++}$ if and only if*

$$\big((I + A_G)\big(t - \tfrac{1}{2}\big) - E \big) \notin \mathcal{COP}.$$

Proof. For any $\lambda \geq \alpha_G$ we have that $\big((I + A_G)\lambda - E \big) \in \mathcal{COP}$. This can be seen from the fact that $\big((I + A_G)\alpha_G - E \big) \in \mathcal{COP}$ and $(I + A_G)$ is nonnegative (so also copositive). Therefore

$$\begin{aligned} \big((I + A_G)\big(t - \tfrac{1}{2}\big) - E \big) \notin \mathcal{COP} &\Leftrightarrow \alpha_G > t - \tfrac{1}{2} \\ &\Leftrightarrow \alpha_G \geq \left\lceil t - \tfrac{1}{2} \right\rceil = t \\ &\Leftrightarrow \text{There is a stable set of size } t. \end{aligned}$$

□

This immediately allows to state the following theorem, confirming again the complexity result by Murty and Kabadi for the copositive cone.

Theorem 2.10. *The stable set problem is Turing reducible to the strong membership problem for the copositive cone with a many-one reduction, and thus the strong membership problem for the copositive cone is NP-hard.*

Proof. This comes from Lemma 2.9, and from noting that the encoding length of $\big((I + A_G)\big(t - \tfrac{1}{2}\big) - E \big)$ is polynomial in the encoding length of the stable set problem. □

In order to extend this to WMEM for the copositive cone we provide the following lemma.

Lemma 2.11. *Let $\varepsilon, \lambda \in \mathbb{R}_+$ such that $\varepsilon \leq 1$ and $\lambda \geq (1-\varepsilon)\alpha_G + \varepsilon(n+1)$. Now define $Z_\lambda := (I + A_G)\lambda - E$. Then we have that $Z_\lambda \in \mathrm{S}(\mathcal{COP}, -\varepsilon)$, or equivalently $\mathrm{S}(Z_\lambda, \varepsilon) \subseteq \mathcal{COP}$.*

Proof. We consider an arbitrary $Y \in \mathrm{S}(Z_\lambda, \varepsilon)$. There exists $V \in \mathbb{S}^n$ such that $\|V\| = 1$ and $Y = Z_\lambda + \varepsilon V$. Now, for all $\mathbf{x} \geq 0$ such that $\|\mathbf{x}\| = 1$ we have

$$\begin{aligned}
\mathbf{x}^\mathsf{T} Y \mathbf{x} &= \mathbf{x}^\mathsf{T}\big((I + A_G)\lambda - E + \varepsilon V\big)\mathbf{x} \\
&= (1-\varepsilon)\mathbf{x}^\mathsf{T}\big((I + A_G)\alpha_G - E\big)\mathbf{x} - \varepsilon\mathbf{x}^\mathsf{T} E\mathbf{x} + \varepsilon\mathbf{x}^\mathsf{T} V\mathbf{x} \\
&\qquad + (\lambda - (1-\varepsilon)\alpha_G)\big(\mathbf{x}^\mathsf{T} I\mathbf{x} + \mathbf{x}^\mathsf{T} A_G\mathbf{x}\big) \\
&\geq 0 - \varepsilon n - \varepsilon + \varepsilon(n+1)(1+0) = 0.
\end{aligned}$$

Therefore $Y \in \mathcal{COP}$, completing the proof. □

We can now state the following lemma and theorem concerning WMEM for the copositive cone.

Lemma 2.12. *Let $G = (V, E)$ be a graph with $|V| = n \in \mathbb{Z}_{++}$ and let*

$$\begin{aligned}
Y &= (I + A_G)\left(t - \tfrac{1}{2}\right) - E, \\
\delta &= 1/(2n+1), \\
K &= \mathcal{COP},
\end{aligned}$$

where $n \geq t \in \mathbb{Z}_{++}$. Then considering the WMEM for these parameters we have that

1. *$Y \in \mathrm{S}(K, \delta)$ would imply that the graph G does not contain a stable set of size t.*
2. *$Y \notin \mathrm{S}(K, -\delta)$ would imply that the graph G does contain a stable set of size t.*

Proof. We shall prove these results separately.

1. Suppose that $Y \in \mathrm{S}(\mathcal{COP}, \delta)$. There must exist $Z \in \mathrm{S}(0, 1)$ such that $(Y + \delta Z) \in \mathcal{COP}$.

   From Lemma 2.11 (setting $\varepsilon = 1$ and $\lambda = n + 1$) we have that

   $$((I + A_G)(n+1) - E - Z) \in \mathcal{COP}.$$

   As the copositive cone is convex the following matrix must again be copositive,

   $$\begin{aligned}
   &\tfrac{\delta}{1+\delta}\Big((I + A_G)(n+1) - E - Z\Big) + \left(1 - \tfrac{\delta}{1+\delta}\right)\Big((I + A_G)\big(t - \tfrac{1}{2}\big) - E + \delta Z\Big) \\
   &\qquad = (I + A_G)\left(t - \frac{2t-1}{4(n+1)}\right) - E.
   \end{aligned}$$

From this we see that $\alpha_G \leq \left\lfloor t - \frac{2t-1}{4(n+1)} \right\rfloor = t - 1$.

Therefore the graph does not contain a stable set of size t.

2. Suppose that $Y \notin \mathrm{S}\left(\mathcal{COP}, -\delta\right)$.

   From Lemma 2.11 we have that $t - \frac{1}{2} < (1-\delta)\alpha_G + \delta(n+1) \leq \alpha_G + \delta n$.

   Therefore $\alpha_G \geq \left\lceil t - \frac{1}{2} - \frac{n}{2n+1} \right\rceil = \left\lceil t - \frac{4n+1}{4n+2} \right\rceil = t$, and so the graph does contain a stable set of size t.

□

Theorem 2.13. *The stable set problem is Turing reducible to the weak membership problem for the copositive cone with a many-one reduction, and thus the weak membership problem for the copositive cone is NP-hard.*

Proof. This comes from Lemma 2.9 and noting that the encoding lengths of Y and δ from this lemma are polynomial in the encoding length of the stable set problem. □

2.3 The Completely Positive Cone

In this section we consider the weak membership problem for the completely positive cone. In order to do this, rather than reformulating to a WMEM problem as we did for the copositive case, this time we reformulate to a WVAL problem, using the following quintuple.

Lemma 2.14. *Consider a graph G with n vertices, let $K = \{X \in \mathcal{CP} \mid \langle I + A_G, X\rangle \leq 1\}$ and furthermore set*

$$\begin{aligned} N &= \tfrac{1}{2}n(n+1), \\ R &= 1, \\ r &= \frac{1}{4n^2}, \\ A_0 &= \tfrac{1}{2n}I + \tfrac{1}{4n^2}E, \end{aligned}$$

then the quintuple $(K; N, R, r, A_0)$ is an A_0-centered convex body as defined in Definition 2.5.

Proof. First we will show that $K \subseteq \mathrm{S}(0, R)$. Note that $(I + A_G)$ is in the interior of the copositive cone, therefore,

$$\begin{aligned}
&\max\{\|X\| \mid X \in K\} \\
&\quad = \max\left\{\|X\| \;\middle|\; X \in \operatorname{conv}\Big(\{0\} \cup \{\mathbf{b}\mathbf{b}^\mathsf{T} \mid \mathbf{b} \in \mathbb{R}^n_+,\ \mathbf{b}^\mathsf{T}(I + A_G)\mathbf{b} = 1\}\Big)\right\} \\
&\quad = \max\left\{\|\mathbf{b}\mathbf{b}^\mathsf{T}\| \;\middle|\; \mathbf{b} \in \mathbb{R}^n_+,\ \mathbf{b}^\mathsf{T}(I + A_G)\mathbf{b} = 1\right\} \\
&\quad \leq 1 = R.
\end{aligned}$$

Next we will show that $S(A_0, r) \subseteq K$. Let $X \in \mathrm{S}(A_0, r)$, then we get the following entrywise inequalities for X,

$$0 \leq \frac{1}{2n}I = A_0 - Er \leq X \leq A_0 + Er = \frac{1}{2n}I + \frac{1}{2n^2}E. \tag{2.4}$$

Hence it now follows that

$$\langle I + A_G, X\rangle \leq \langle E, X\rangle \leq \frac{1}{2n}\langle E, I\rangle + \frac{1}{2n^2}\langle E, E\rangle = \frac{n}{2n} + \frac{n^2}{2n^2} = 1.$$

What is left to show now is that $X \in \mathcal{CP}$. To do this we use a result from Kaykobad [Kay87] that says that if we have an entrywise nonnegative matrix $Y \in \mathbb{S}^n$, which is diagonally dominant, i.e. $(Y)_{ii} \geq \sum_{j \neq i}(Y)_{ij}$ for all $i = 1, \ldots, n$, then $Y \in \mathcal{CP}$.

From (2.4), we note that for all i, j such that $i \neq j$, we have $0 \leq X_{ij} \leq \frac{1}{2n^2}$ and $\frac{1}{2n} \leq X_{ii}$. The result then immediately follows. □

We can now state the following theorem from which the NP-hardness of MEM for the completely positive cone will follow.

Theorem 2.15. *Consider a graph G with n vertices, let $t \in \mathbb{Z}_{++}$, and let $(K; N, R, r, A_0)$ be the A_0-centered convex body as described in Lemma 2.14. Then set*

$$\begin{aligned}
C &= E, \\
\gamma &= t - \tfrac{1}{2}, \\
\varepsilon &= \frac{1}{16n^2t}.
\end{aligned}$$

Considering the WVAL for these parameters we have that

1. *$\langle C, X\rangle \leq \gamma + \varepsilon$ for all $X \in \mathrm{S}(K, -\varepsilon)$ would imply that the graph does not contain a stable set of size t.*

2. $\exists Y \in \mathrm{S}(K, \varepsilon)$ such that $\langle C, Y\rangle \geq \gamma - \varepsilon$ would imply that the graph does contain a stable set of size t.

From this we then have that the stable set problem is Turing reducible to WMEM for K.

Proof. Using Lemma 2.6 we look at what the results of WVAL for our choice of parameters would mean. A major point in the implications is that $\alpha_G \in \mathbb{Z}_{++}$ and we recall from (2.3) that

$$\alpha_G = \max\{\langle C, X\rangle \mid X \in K\}.$$

1. Let $\langle C, X\rangle \leq \gamma + \varepsilon$, for all $X \in \mathrm{S}(K, -\varepsilon)$. Then for all $Z \in K$ we have that

$$\begin{aligned}
\langle C, Z\rangle &\leq \left(\gamma + \varepsilon - \frac{\varepsilon}{r}\langle C, A_0\rangle\right) \Big/ \left(1 - \frac{\varepsilon}{r}\right) \\
&= \left(t - \frac{1}{2} + \frac{1}{16n^2t} - \frac{4n^2}{16n^2t}\left\langle E, \frac{1}{2n}I + \frac{1}{4n^2}E\right\rangle\right) \Big/ \left(1 - \frac{4n^2}{16n^2t}\right) \\
&= \left(\left(t - \tfrac{1}{4}\right)(16n^2t - 4n^2) - 4n^2 + 1\right) \big/ \left(16n^2t - 4n^2\right) \\
&= t - \tfrac{1}{4} - \frac{4n^2 - 1}{4n^2(4t-1)} \leq t - \tfrac{1}{4}.
\end{aligned}$$

   Therefore $\alpha_G \leq \lfloor t - \frac{1}{4}\rfloor = t - 1$ and so the graph does not contain a stable set of size t.

2. Assume $\exists Y \in \mathrm{S}(K, \varepsilon)$ such that $\langle C, Y\rangle \geq \gamma - \varepsilon$. Then again by Lemma 2.6, $\exists Z \in K$ such that

$$\begin{aligned}
\langle C, Z\rangle &\geq \gamma - (1 + \|C\|)\varepsilon \\
&= t - \frac{5}{8} + \frac{2n^2t - n - 1}{16n^2t} \\
&\geq t - \frac{5}{8}.
\end{aligned}$$

   Therefore $\alpha_G \geq \lceil t - \frac{5}{8}\rceil = t$ and so the graph does contain a stable set of size t.

It can now be seen that the stable set problem is Turing reducible to WMEM for K by using Theorem 2.8 and noting that the encoding lengths of $(K; N, R, r, A_0)$, C, γ and ε are polynomial in the encoding length of the stable set problem. □

We now get the following result on MEM for the completely positive cone.

Theorem 2.16. *The strong membership problem for the completely positive cone is NP-hard.*

Proof. From Theorem 2.15, we have that the stable set problem is Turing reducible to WMEM for K. This is in turn many-one reducible to MEM for the completely positive cone, as the encoding length of $(I + A_G)$ is polynomial in the encoding length of a matrix in $\mathbb{S}^n$. $\square$

In order to extend this to WMEM for the completely positive cone we need a way of solving WMEM for K from Lemma 2.14 given a weak membership oracle for the completely positive cone.

Lemma 2.17. *We consider K from Lemma 2.14, define $\mathcal{H} := \{X \in \mathbb{S} \mid \langle I + A_G, X\rangle \leq 1\}$ and let $\delta \in \mathbb{Q}_{++}$. Then we have that*

1. $\mathrm{S}\,(K, -\delta) = \mathrm{S}\,(\mathcal{H}, -\delta) \cap \mathrm{S}\,(\mathcal{CP}, -\delta) \subseteq \mathcal{H} \cap \mathrm{S}\left(\mathcal{CP}, -\delta/(1 + n^2)\right)$,
2. $\mathrm{S}\,(K, \delta) \supseteq \mathcal{H} \cap \mathrm{S}\left(\mathcal{CP}, \delta/(1 + n^2)\right)$.

From this we then have that WMEM for K is Turing reducible to WMEM for the completely positive cone.

Proof. We consider each of these parts separately.

1. This comes from noting that

$$
\begin{aligned}
\mathrm{S}\,(K, -\delta) &= \{X \in \mathbb{S} \mid \mathrm{S}\,(X, \delta) \subseteq \mathcal{H} \cap \mathcal{CP}\} \\
&= \{X \in \mathbb{S} \mid \mathrm{S}\,(X, \delta) \subseteq \mathcal{H}\} \cap \{X \in \mathbb{S} \mid \mathrm{S}\,(X, \delta) \subseteq \mathcal{CP}\} \\
&= \mathrm{S}\,(\mathcal{H}, -\delta) \cap \mathrm{S}\,(\mathcal{CP}, -\delta)\,, \\
\mathrm{S}\,(\mathcal{H}, -\delta) &\subseteq \mathcal{H}, \\
\mathrm{S}\,(\mathcal{CP}, -\delta) &\subseteq \mathrm{S}\left(\mathcal{CP}, -\delta/(1 + n^2)\right).
\end{aligned}
$$

2. Consider an arbitrary $X \in \mathcal{H} \cap \mathrm{S}\left(\mathcal{CP}, \delta/(1 + n^2)\right)$ and let $\varepsilon = \delta/(1+n^2)$. Then there exists $Y \in \mathcal{CP}$ such that $\|X - Y\| \leq \varepsilon$ and we have that

$$\langle I + A_G, Y\rangle \leq \langle I + A_G, X\rangle + \varepsilon\|I + A_G\| \leq 1 + \varepsilon n^2,$$

$$
\begin{aligned}
&\|Y\| \leq \max\left\{\|U\| \;\middle|\; U \in \mathcal{CP},\ \langle I + A_G, U\rangle \leq 1 + \varepsilon n^2\right\} \\
&= \max\left\{\|U\| \;\middle|\; U \in \operatorname{conv}\left(\{0\} \cup \{\mathbf{b}\mathbf{b}^\mathsf{T} \mid \mathbf{b} \in \mathbb{R}^n_+,\ \mathbf{b}^\mathsf{T}(I + A_G)\mathbf{b} = 1 + \varepsilon n^2\}\right)\right\} \\
&\leq 1 + \varepsilon n^2.
\end{aligned}
$$

We now let $Z = \frac{1}{1+\varepsilon n^2}Y$. We have that $Z \in \mathcal{CP}$ and $\langle I + A_G, Z\rangle \leq 1$. Therefore $Z \in \mathcal{H} \cap \mathcal{CP}$. We finish the proof by noting that

$$\begin{aligned}
\|X - Z\| &= \left\| X - \frac{1}{1+\varepsilon n^2}Y \right\| \\
&= \left\| X - Y + \frac{\varepsilon n^2}{1+\varepsilon n^2}Y \right\| \\
&\leq \|X - Y\| + \frac{\varepsilon n^2}{1+\varepsilon n^2}\|Y\| \\
&\leq \varepsilon(1+n^2) \\
&= \delta,
\end{aligned}$$

and hence $X \in \mathrm{S}\left(\mathcal{H} \cap \mathcal{CP}, \delta\right) = \mathrm{S}\left(K, \delta\right)$.

We then have that WMEM for K is Turing reducible to WMEM for the completely positive cone by noting that the encoding lengths of $(I + A_G)$ and $\delta/(1+n^2)$ are polynomial in the encoding lengths of the input δ and a matrix in $\mathbb{S}^n$. □

We now get the following result, which is the main result of this paper,

Theorem 2.18. *Both the weak and strong membership problems for the completely positive cone are NP-hard.*

Proof. From Theorem 2.15, and Lemma 2.17, it follows that the stable set problem is Turing reducible to WMEM for the completely positive cone, and thus this problem is NP-hard. To show that MEM for the completely positive cone is also NP-hard we can either use Theorem 2.16 or a many-one reduction to WMEM for the completely positive cone. □

This finally establishes that the strong and weak membership problems for the copositive as well as the completely positive cone are in NP-hard. As we mentioned in the Introduction of this thesis whether or not membership for the completely positive cone is in NP is still an open question at the moment of writing. Another interesting question that can be asked considering the results of this chapter is whether or not this type of complexity result holds in general. More specifically we define the following open problem.

Open Problem 2.19. *Let $K \in \mathbb{R}^{n\times n}$ and suppose that the (weak) membership problem for K is NP-hard. Does this then imply that the (weak) membership problem for K^* is also NP-hard.*

Chapter 3

Irreducible elements of the copositive cone [1]

[1]Published as [DDGH13a] Peter J.C. Dickinson, Mirjam Dür, Luuk Gijben, and Roland Hildebrand. Irreducible elements of the copositive cone. *Linear Algebra and its Applications*, 439(6):1605-1626, 2013.

In this chapter, we study the structure of $\mathcal{COP}^n$ in relation to the cones $\mathcal{S}_+^n$ and $\mathcal{N}^n$. We recall from the introduction that if $A \in \mathcal{S}_+^n$ or $A \in \mathcal{N}^n$, then A must be copositive, that is $\mathcal{S}_+^n + \mathcal{N}^n \subseteq \mathcal{COP}^n$. Moreover recall that $\mathcal{S}_+^n + \mathcal{N}^n = \mathcal{COP}^n$ if and only if $n \leq 4$.

From an optimization viewpoint, optimizing over the cone $\mathcal{S}_+^n + \mathcal{N}^n$ is computationally easy as we saw in the Introduction and can be done by standard algorithms like interior point methods, whereas optimizing over $\mathcal{COP}^n$ is hard as we have just seen in the previous chapter. Therefore, the cone $\mathcal{COP}^n$ for $n \geq 5$ is of special interest. Hence, in this chapter we investigate the structure of $\mathcal{COP}^n$ for $n \geq 5$, where we have $\mathcal{COP}^n \neq \mathcal{S}_+^n + \mathcal{N}^n$. Matrices in $\mathcal{COP}^n \setminus (\mathcal{S}_+^n + \mathcal{N}^n)$ have been studied by Johnson and Reams [JR08] who baptized those matrices *exceptional matrices.* As noted before, the Horn matrix (1.10) is in $\mathcal{COP}^n \setminus (\mathcal{S}_+^n + \mathcal{N}^n)$ and therefore is an example of an exceptional matrix.

There are still a number of open problems concerning the cone $\mathcal{COP}^n$, the characterization of its extreme rays as discussed in the introduction is probably the biggest one. In this chapter we will consider a weaker concept than extremality, namely that of irreducibility with respect to the cone of nonnegative matrices. This property has been utilized and recognized as being more convenient than extremality already in the early work on copositive matrices [Dia62],[HN63],[Bau67]. In particular, it was studied by Baumert [Bau65] who gave a characterization of irreducible matrices. In this chapter, we demonstrate that Baumert's characterization is incorrect and give a correct version of his theorem which establishes a necessary and sufficient condition for a copositive matrix to be irreducible.

For the case of 5×5 copositive matrices we give a complete characterization of all irreducible matrices. We show that those irreducible matrices in $\mathcal{COP}^5$ which are not positive semidefinite can be parameterized in a semi-trigonometric way. Finally, we prove that every 5×5 copositive matrix which is not the sum of a nonnegative and a semidefinite matrix can be expressed as the sum of a nonnegative and a single irreducible matrix. This last result can be seen as a dual statement to [BAD09, Corollary 2], where it was shown that in the 5×5 case any doubly nonnegative matrix which is not completely positive (such matrices are called "bad matrices" in [BAD09]) can be written as the sum of a completely positive matrix and a single ”extremely bad” matrix (i.e., a matrix which is extremal for the doubly nonnegative cone, but not completely positive).

3.1 Notation

During this chapter we let $\mathbf{e}_i$ be the unit vector with the ith element equal to one and all other elements equal to 0. For simplicity we shall also denote $\mathbf{e}_{ij} = \mathbf{e}_i + \mathbf{e}_j$ for $i \neq j$.
For a vector $\mathbf{u} \in \mathbb{R}^n_+$ we define its support as

$$\text{supp}(\mathbf{u}) := \{i \in \{1, \dots, n\} \mid u_i > 0\}.$$

For a set $\mathcal{M} \subseteq \mathbb{R}^n_+$ we shall define its support as

$$\text{supp}(\mathcal{M}) := \{\text{supp}(\mathbf{u}) \mid \mathbf{u} \in \mathcal{M}\}.$$

For a subset $\mathcal{I} \subseteq \{1, \dots, n\}$ we denote by $A_{\mathcal{I}}$ the principal submatrix of A whose elements have row and column indices in $\mathcal{I}$, i.e. $A_{\mathcal{I}} := (a_{ij})_{i,j \in \mathcal{I}}$. Similarly for a vector $\boldsymbol{v} \in \mathbb{R}^n$ we denote the subvector $\mathbf{u}_{\mathcal{I}} := (u_i)_{i \in \mathcal{I}}$.
For $i, j = 1, \dots, n$, we denote the following generators of the extreme rays of the nonnegative cone $\mathcal{N}^n$ by

$$E_{\{ij\}} := \begin{cases} \mathbf{e}_i \mathbf{e}_i^\mathsf{T} & \text{if } i = j \\ \mathbf{e}_i \mathbf{e}_j^\mathsf{T} + \mathbf{e}_j \mathbf{e}_i^\mathsf{T} & \text{otherwise.} \end{cases}$$

This notation is not to be confused with E, the matrix of all ones, or the matrices E^{ij} that were defined before, and who are in a sense the non-symmetric versions of the matrices $E_{\{ij\}}$. We call a nonzero vector $\mathbf{u} \in \mathbb{R}^n_+$ a *zero* of a copositive matrix $A \in \mathcal{COP}^n$ if $\mathbf{u}^\mathsf{T} A \mathbf{u} = \mathbf{0}$ We denote the set of zeros of A by

$$\mathcal{V}^A := \{\mathbf{u} \in \mathbb{R}^n_+ \setminus \{\mathbf{0}\} \mid \mathbf{u}^\mathsf{T} A \mathbf{u} = 0\}.$$

This notation is similar to that used by Dickinson in [Dic10], except for the fact that we will exclude the vector $\mathbf{0}$ from the set.

Let $\text{Aut}(\mathbb{R}^n_+)$ be the automorphism group of the nonnegative orthant. It is generated by all $n \times n$ permutation matrices and by all $n \times n$ diagonal matrices with strictly positive diagonal elements. This group generates a group $\mathcal{G}_n$ of automorphisms of $\mathcal{COP}^n$ by $A \mapsto AXA^\mathsf{T}$, $A \in \text{Aut}(\mathbb{R}^n_+)$. Whenever we speak of orbits of elements in $\mathcal{COP}^n$, we mean orbits with respect to the action of $\mathcal{G}_n$.

Definition 3.1. For a matrix $A \in \mathcal{COP}^n$ and a set $\mathcal{M}$ contained in the space of symmetric matrices, we say that A is *$\mathcal{M}$-irreducible* if there do not exist $\gamma > 0$ and $M \in \mathcal{M} \setminus \{0\}$ such that $A - \gamma M \in \mathcal{COP}^n$.

Note that this definition differs from the concept of an irreducible matrix that is normally used in matrix theory. We will synonymously use the expressions of being $\mathcal{M}$-irreducible and being irreducible with respect to $\mathcal{M}$. For simplicity we speak about irreducibility with respect to M when $\mathcal{M} = \{M\}$. During this chapter, we shall be concerned with the cases

$$\mathcal{M} = \mathcal{N}^n, \quad \mathcal{M} = \widetilde{\mathcal{N}}^n, \quad \text{and} \quad \mathcal{M} = \{E_{\{ij\}}\}.$$

where $\widetilde{\mathcal{N}}^n := \{N \in \mathcal{N}^n \mid \operatorname{diag}(N) = \mathbf{0}\}$. The $A^*(n;0)$-property defined in [Dia62, p.17] or, equivalently, the $A^*(n)$-property[2] defined in [Bau66, Def. 2.1] are then equivalent to being irreducible with respect to $\mathcal{N}^n$.

It is easy to see that an irreducible matrix necessarily is in the boundary of $\mathcal{COP}^n$. Also note that if a nonzero matrix A is on an extreme ray of $\mathcal{COP}^n$, then A must be $\mathcal{N}^n$-irreducible: indeed, assume the contrary. Then there exist $\gamma > 0$ and $0 \neq N \in \mathcal{N}^n$ such that $A - \gamma N =: B \in \mathcal{COP}^n$. But then $A = B + \gamma N$, contradicting extremality.

Observe that being $\widetilde{\mathcal{N}}^n$-irreducible is a weaker condition than being $\mathcal{N}^n$-irreducible. It is also trivial to see that $A \in \mathcal{COP}^n$ is irreducible with respect to $\mathcal{N}^n$ if and only if it is irreducible with respect to $E_{\{ij\}}$ for all $i, j = 1, \ldots, n$, whilst A is irreducible with respect to $\widetilde{\mathcal{N}}^n$ if and only if it is irreducible with respect to $E_{\{ij\}}$ for all $i \neq j$.

3.2 Irreducible copositive matrices

In this section we shall show that the following theorem from Baumert's thesis [Bau65] about irreducibility with respect to $E_{\{ij\}}$ is incorrect, and we shall give the corrected version of this theorem.

Assertion 3.2 (Incorrect Theorem 3.3 of [Bau65])**.** *For $i, j = 1, \ldots, n$ we have that a matrix $A \in \mathcal{COP}^n$ is irreducible with respect to $E_{\{ij\}}$ if and only if there exists a vector $\mathbf{u} \in \mathcal{V}^A$ such that $u_i u_j > 0$.*

It is trivial to see that the reverse implication holds, as for such a $\mathbf{u}$ and any $\gamma > 0$ we have that $\mathbf{u}^\mathsf{T}(A - \gamma E_{\{ij\}})\mathbf{u} = -2\gamma u_i u_j < 0$. However the following

[2] This notation is also used in [Dia62], but without definition.

matrix is a counterexample to the forward implication:

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{pmatrix} \tag{3.1}$$

In order to show this we need the following lemma.

Lemma 3.3. *Let $A \in \mathcal{S}_+^n$. Then $\mathbf{u} \in \mathbb{R}^n$ is a zero of A, that is $\mathbf{u}^\mathsf{T} A\mathbf{u} = 0$, if and only if $A\mathbf{u} = \mathbf{0}$.*

Proof. Let $A \in \mathcal{S}_+^n$ and suppose that $\mathbf{u} \in \mathbb{R}^n$ is such that $\mathbf{u}^\mathsf{T} A\mathbf{u} = 0$. Then because $A \in \mathcal{S}_+^n$ we can write A $= BB^\mathsf{T}$ and hence

$$0 = \mathbf{u}^\mathsf{T} A\mathbf{u} = \mathbf{u}^\mathsf{T} BB^\mathsf{T}\mathbf{u} = (B^\mathsf{T}\mathbf{u})^\mathsf{T}(B^\mathsf{T}\mathbf{u}) = \mathbf{y}^\mathsf{T}\mathbf{y} = \sum_{i=1}^{n}(y_i)^2.$$

This can obviously only happen if $y_i = 0$ for all $i = 1, \ldots, n$, which means that we have $A\mathbf{u} = B(B^\mathsf{T}\mathbf{u}) = \mathbf{0}$.

On the other hand if $A\mathbf{u} = \mathbf{0}$ then obviously $\mathbf{u}^\mathsf{T} A\mathbf{u} = 0$. □

Now because A, as defined by (3.1), is positive semidefinite and because we have that $\mathbf{x}^\mathsf{T} A\mathbf{x} = 0$ if and only if $A\mathbf{x} = \mathbf{0}$ we get that $\mathcal{V}^A = \{\lambda\mathbf{e}_{23} \mid \lambda > 0\}$. Therefore, according to Baumert's theorem, A should be irreducible *only* with respect to $E_{\{23\}}$, and hence A should *not* be irreducible with respect to, say, $E_{\{12\}}$. In other words, there should exist a $\gamma > 0$ such that $A - \gamma E_{\{12\}}$ is copositive. However, for any $\gamma > 0$ we have

$$\begin{pmatrix} \gamma \\ 1 \\ 1 \end{pmatrix}^\mathsf{T} \begin{pmatrix} 1 & -\gamma & 0 \\ -\gamma & 1 & -1 \\ 0 & -1 & 1 \end{pmatrix} \begin{pmatrix} \gamma \\ 1 \\ 1 \end{pmatrix} = -\gamma^2 < 0,$$

which demonstrates that A is also irreducible with respect to $E_{\{12\}}$, contradicting Baumert's theorem.

As [Bau65, Theorem 3.3] is incorrect[3], also other results from Baumert's

[3]The proof of Theorem 3.3 of [Bau65] contains two errors, each of which leads to counterexamples. One error is the omission of the multiples of $E_{\{11\}}$ in the classification making up the inductive base on page 12 of Baumert's thesis [Bau65]. (Note that for the 2×2 case, the matrix $E_{\{11\}}$ also provides a counterexample to Baumert's assertion.) The other error, being located in the induction step for case 3 on page 14, is more subtle. Generally, in order for an induction step to be applicable, the problem in question has to be reduced to an instance of the same problem of strictly smaller size. In the proof this is to be ensured by the fact that the copositive form Q_2 depends on a strictly smaller number of variables than the original form Q, in particular, because Q_2 does no more depend on the variable x_2. However, since irreducibility with respect to $E_{\{12\}}$ is studied, dependence on x_2 has nevertheless to be formally included. Therefore, the induction step is applicable only if either the support of the zero $\mathbf{u}$ is strictly bigger than $\{2\}$, or the support of $Q\mathbf{u}$ is nonempty. Both conditions may simultaneously fail to be satisfied, however.

thesis [Bau65] which are proved using this result must be treated with care. Fortunately this error has not entered into Baumert's papers [Bau66, Bau67], which as far as we can tell are correct. In our paper we do use results from Baumert's papers, and to make sure that this error could not have had an effect, we made sure that they were correct from first principles.

For easy reference in the remainder of the paper, we reference the property used in Baumert's incorrect theorem as follows:

Property 3.4. *For each pair i, j of indices such that $i \neq j$, there exists a zero $\mathbf{u}$ of A such that $u_i u_j > 0$.*

Our next aim is to present a corrected version of this theorem. Before doing so, however, we first discuss a couple of properties related to the set of zeros.

Lemma 3.5 (p. 200 of [Bau66])**.** *Let $A \in \mathcal{COP}^n$ and $\mathbf{u} \in \mathcal{V}^A$. Then $A\mathbf{u} \geq \mathbf{0}$.*

Lemma 3.6. *Let $A \in \mathcal{COP}^n$ and $\mathbf{u} \in \mathcal{V}^A$. Then the principal submatrix $A_{\mathrm{supp}(\mathbf{u})}$ is positive semidefinite.*

Proof. Clearly, $\mathbf{u}$ is a zero of A if and only if $\boldsymbol{u}_{\mathrm{supp}(\mathbf{u})}$ is a zero of $A_{\mathrm{supp}(\mathbf{u})}$. Moreover, $\mathbf{u}_{\mathrm{supp}(\mathbf{u})} > \mathbf{0}$. Therefore, by [Dia62, Lemma 1], we have that

$$(\mathbf{u}_{\mathrm{supp}(\mathbf{u})})^\mathsf{T} \, A_{\mathrm{supp}(\mathbf{u})} \, (\mathbf{u}_{\mathrm{supp}(\mathbf{u})}) \;=\; \mathbf{0}$$

implies that $A_{\mathrm{supp}(\mathbf{u})}$ is positive semidefinite. □

Lemma 3.7. *Let $A \in \mathcal{COP}^n$ and $\mathbf{u}, \mathbf{v} \in \mathcal{V}^A$ such that $\mathrm{supp}(\mathbf{u}) \subseteq \mathrm{supp}(\mathbf{v})$. Then we have that $(A\mathbf{u})_i = 0$ for all $i \in \mathrm{supp}(\mathbf{v})$.*

Proof. This comes trivially from the previous lemma and recalling that for arbitrary $B \in \mathcal{S}_+^m$ and $\mathbf{v} \in \mathbb{R}^m$ we have that $\mathbf{v}^\mathsf{T} B\mathbf{v} = 0$ if and only if $B\mathbf{v} = \mathbf{0}$. □

Note that applying the previous lemma for $\mathbf{u} = \mathbf{v}$ gives us that $A \in \mathcal{COP}^n$ and $\mathbf{u} \in \mathcal{V}^A$ imply that $(A\mathbf{u})_i = 0$ for all $i \in \mathrm{supp}(\mathbf{u})$.

We now present the main theorem of this section, which is the corrected version of Baumert's theorem.

Theorem 3.8. *Let $A \in \mathcal{COP}^n$, $n \geq 2$, and let $1 \leq i, j \leq n$. Then the following conditions are equivalent.*

(i) A is irreducible with respect to $E_{\{ij\}}$,

(ii) there exists $\mathbf{u} \in \mathcal{V}^A$ such that $(A\mathbf{u})_i = (A\mathbf{u})_j = 0$ and $u_i + u_j > 0$.

Proof. The special case when $i = j$ is proven in [Bau66, Theorem 3.4], so from now on we shall consider $i \neq j$. For $\varepsilon > 0$, we will abbreviate $A_\varepsilon := A - \varepsilon E_{\{ij\}}$.

We first show (ii) $\Rightarrow$ (i). Assume that there exists such a $\mathbf{u} \in \mathcal{V}^A$. Fix $\varepsilon > 0$ and let $\delta > 0$. Then $\mathbf{u} + \delta \mathbf{e}_{ij} \geq 0$ and

$$\begin{aligned}(\mathbf{u} + \delta\mathbf{e}_{ij})^\mathsf{T} A_\varepsilon (\mathbf{u} + \delta\mathbf{e}_{ij}) &= \mathbf{u}^\mathsf{T} A\mathbf{u} + 2\delta\mathbf{e}_{ij}^\mathsf{T} A\mathbf{u} + \delta^2(a_{ii} + 2a_{ij} + a_{jj}) - 2\varepsilon(u_i + \delta)(u_j + \delta) \\ &= -2\varepsilon u_i u_j - 2\varepsilon\delta(u_i + u_j) + \mathcal{O}(\delta^2).\end{aligned}$$

Since $u_i u_j \geq 0$ and $u_i + u_j > 0$, this term is negative for $\delta > 0$ small enough. Hence A_ε is not copositive and A satisfies condition (i).

Let us now show (i) $\Rightarrow$ (ii) by induction over n. For $n = 2$, it can be seen that copositive matrices satisfying condition (i) are of the form

$$\begin{pmatrix} a^2 & -ab \\ -ab & b^2 \end{pmatrix} \quad \text{with } a, b \geq 0.$$

If $a = b = 0$, then any $\boldsymbol{u} \in \mathbb{R}^n_+ \setminus \{\mathbf{0}\}$ satisfies condition (ii). Alternatively, if $a + b > 0$, then the zero $\mathbf{u} = (b, a)^\mathsf{T}$ satisfies condition (ii).

Now assume that the assertion holds for $\hat{n} < n$, and let $A \in \mathcal{COP}^n$ satisfy condition (i). For every $\varepsilon > 0$, consider the optimization problem

$$\min\left\{\tfrac{1}{2}\mathbf{v}^\mathsf{T} A_\varepsilon \mathbf{v} \mid \mathbf{v} \in \mathbb{R}^n_+,\ \mathbf{e}^\mathsf{T}\mathbf{v} = 1\right\}. \tag{3.2}$$

By condition (i) the optimal value of this problem is negative, and it is attained by compactness of the feasible set. Let $\boldsymbol{v}$ be a minimizer of the problem. Having only linear constraints, the problem fulfills a constraint qualification, and therefore it follows from the Karush-Kuhn-Tucker optimality conditions that there exist Lagrange multipliers $\boldsymbol{\lambda} \in \mathbb{R}^n_+$ and $\mu \in \mathbb{R}$ such that $\mathbf{v}^\mathsf{T}\boldsymbol{\lambda} = 0$ and $A_\varepsilon\mathbf{v} - \boldsymbol{\lambda} + \mu\mathbf{e} = 0$, or equivalently,

$$A_\varepsilon\mathbf{v} = \boldsymbol{\lambda} - \mu\mathbf{e}. \tag{3.3}$$

Multiplying with $\mathbf{v}^\mathsf{T}$ and observing $\mathbf{v}^\mathsf{T}\boldsymbol{\lambda} = 0$ and $\mathbf{v}^\mathsf{T}\mathbf{e} = 1$, we obtain $-\mu = \mathbf{v}^\mathsf{T} A_\varepsilon \mathbf{v} < 0$. We also have $\mathbf{v}^\mathsf{T} A\mathbf{v} - 2\varepsilon v_i v_j = \mathbf{v}^\mathsf{T} A_\varepsilon\mathbf{v} < 0$, which by $\mathbf{v}^\mathsf{T} A\mathbf{v} \geq 0$ yields $v_i > 0$, $v_j > 0$. Therefore $\lambda_i = \lambda_j = 0$, and

$$(A_\varepsilon\mathbf{v})_i = (A_\varepsilon\mathbf{v})_j = -\mu < 0. \tag{3.4}$$

Let now $\varepsilon_k \to 0$, let $\mathbf{v}^k \in \mathbb{R}^n_+$ be a minimizer of problem (3.2) for $\varepsilon = \varepsilon_k$, and let $\boldsymbol{\lambda}^k = (\lambda^k_1, \dots, \lambda^k_n)$, μ^k be the corresponding Lagrange multipliers. By possibly choosing a subsequence, we can assume w.l.o.g. that $\mathbf{v}^k \to \mathbf{v}^* \in \mathbb{R}^n_+ \setminus \{\mathbf{0}\}$. Then $0 \geq \lim_{k\to\infty} {\mathbf{v}^k}^\mathsf{T} A_{\varepsilon_k}\mathbf{v}^k = {\mathbf{v}^*}^\mathsf{T} A\mathbf{v}^* \geq 0$, and hence by Lemma 3.5

$A\mathbf{v}^* \geq \mathbf{0}$. Since $\lim_{k\to\infty} A_{\varepsilon_k}\mathbf{v}^k = A\mathbf{v}^* \geq \mathbf{0}$, we get from (3.4) that $(A\mathbf{v}^*)_i = (A\mathbf{v}^*)_j = 0$. If $v_i^* + v_j^* > 0$, then $\mathbf{v}^*$ verifies condition (ii). Hence assume that $v_i^* = v_j^* = 0$ and consider two cases:

Case 1: Suppose there exists an index l such that $(A\mathbf{v}^*)_l > 0$. Then $i \neq l \neq j$. From (3.3) we obtain $(A_{\varepsilon_k}\mathbf{v}^k)_l = \lambda_l^k - \mu^k < \lambda_l^k$ for all k. By $\lim_{k\to\infty}(A_{\varepsilon_k}\mathbf{v}^k)_l = (A\mathbf{v}^*)_l > 0$ we must have $\lambda_l^k > 0$ for k large enough, and hence $(\boldsymbol{\lambda}^k)^\mathsf{T}\mathbf{v}^k = 0$ implies $\mathbf{v}_l^k = 0$ for those k. This implies that

$$0 > (\mathbf{v}^k)^\mathsf{T}(A - \varepsilon_k E_{\{ij\}})\mathbf{v}^k = (\hat{\mathbf{v}}^k)^\mathsf{T}(\hat{A} - \varepsilon_k \hat{E}_{\{ij\}})\hat{\mathbf{v}}^k$$

where $\hat{A}$ (resp. $\hat{E}_{\{ij\}}$) are the principal submatrices of A (resp. $E_{\{ij\}}$) obtained by crossing out row and column l, and $\hat{\mathbf{v}}^k$ is obtained from $\mathbf{v}^k$ by crossing out row l. This means that $\hat{A}$ is irreducible with respect to $\hat{E}_{\{ij\}}$. By the induction hypothesis, there exists a zero $\hat{\mathbf{u}}$ of $\hat{A}$ satisfying condition (ii). The sought zero $\mathbf{u}$ of A can then be obtained by inserting a '0' in $\hat{\mathbf{u}}$ at position l, which concludes the proof for this case.

Case 2: Suppose now that $A\mathbf{v}^* = \mathbf{0}$. Let $\mathcal{I} = \operatorname{supp}(\mathbf{v}^*)$ and let $\mathcal{J} = \{1, \dots, n\} \setminus \mathcal{I}$. Note that $i, j \in \mathcal{J}$ and that $|\mathcal{J}| < n$. By [Bau66, Lemma 3.1] we can represent A as $A = P + C$, where P is positive semidefinite, the principal submatrix $P_\mathcal{I}$ has the same rank as P, and C is a copositive matrix whose principal submatrix $C_\mathcal{J}$ contains all its nonzero elements. The property ${\mathbf{v}^*}^\mathsf{T} A\mathbf{v}^* = 0$ implies $P\mathbf{v}^* = \mathbf{0}$ and ${\mathbf{v}^*}^\mathsf{T} C\mathbf{v}^* = 0$. Moreover, $C\mathbf{v}^* = \mathbf{0}$ by construction. The matrix C is irreducible with respect to $E_{\{ij\}}$, and hence $C_\mathcal{J}$ is irreducible with respect to $(E_{\{ij\}})_\mathcal{J}$. By the induction hypothesis on $C_\mathcal{J}$ there exists a zero with the sought properties which we can augment to obtain a zero $\mathbf{u}$ of C with $(C\mathbf{u})_i = (C\mathbf{u})_j = 0$, $u_i + u_j > 0$, and $u_l = 0$ for all $l \in \mathcal{I}$. By the rank property of P there exists a vector $\tilde{\mathbf{u}}$ such that $\tilde{u}_l = u_l$ for all $l \in \mathcal{J}$ and $P\tilde{\mathbf{u}} = \mathbf{0}$. Since $\tilde{u}_l \geq 0$ for all $l \in \mathcal{J}$, there exists $\alpha \geq 0$ such that $\mathbf{w} := \tilde{\mathbf{u}} + \alpha\mathbf{v}^* \geq \mathbf{0}$. Then we have $P\mathbf{w} = \mathbf{0}$ and $\mathbf{w}^\mathsf{T} A\mathbf{w} = \mathbf{w}^\mathsf{T} C\mathbf{w} = \tilde{\mathbf{u}}^\mathsf{T} C\tilde{\mathbf{u}} = \mathbf{u}^\mathsf{T} C\mathbf{u} = 0$ and $A\mathbf{w} = C\mathbf{w} = C\tilde{\mathbf{u}} = C\mathbf{u}$. It follows that $(A\mathbf{w})_i = (A\mathbf{w})_j = 0$, and moreover $w_i + w_j = u_i + u_j > 0$. Thus A satisfies condition (ii), which completes the proof of the theorem. □

Related to this theorem we define the following.

Definition 3.9. Let $A \in \mathcal{COP}^n$ be irreducible with respect to $E_{\{ij\}}$. If $\mathbf{u} \in \mathcal{V}^A$ satisfies condition (ii) in Theorem 3.8, then we say that irreducibility with respect to $E_{\{ij\}}$ is *associated to* $\mathbf{u}$, or for short, $E_{\{ij\}}$ *is associated to* $\mathbf{u}$.

As a consequence of Theorem 3.8, a matrix $A \in \mathcal{COP}^n$ is irreducible with respect to $\widetilde{\mathcal{N}}^n$ if and only if it satisfies the following property.

Property 3.10. *For each index pair (i,j) with $i < j$, there exists $\mathbf{u} \in \mathcal{V}^A$ such that at least one of the following holds:*

- $u_i u_j > 0$,
- $u_i > 0$ *and* $u_j = (A\mathbf{u})_j = 0$,
- $u_j > 0$ *and* $u_i = (A\mathbf{u})_i = 0$.

Another consequence of Theorem 3.8 is that a matrix $A \in \mathcal{COP}^n$ is irreducible with respect to $\mathcal{N}^n$ if and only if it satisfies Property 3.10 and for all $i = 1, \ldots, n$ there exists $\mathbf{u} \in \mathcal{V}^A$ with $u_i > 0$.

Note that Property 3.10 is weaker than Property 3.4. In the next section we consider explicit examples of matrices $A \in \mathcal{COP}^5$ which satisfy Property 3.10, but not Property 3.4.

3.3 S-matrices

In this section we consider matrices of the form

$$S(\boldsymbol{\theta}) = \begin{pmatrix} 1 & -\cos\theta_1 & \cos(\theta_1+\theta_2) & \cos(\theta_4+\theta_5) & -\cos\theta_5 \\ -\cos\theta_1 & 1 & -\cos\theta_2 & \cos(\theta_2+\theta_3) & \cos(\theta_5+\theta_1) \\ \cos(\theta_1+\theta_2) & -\cos\theta_2 & 1 & -\cos\theta_3 & \cos(\theta_3+\theta_4) \\ \cos(\theta_4+\theta_5) & \cos(\theta_2+\theta_3) & -\cos\theta_3 & 1 & -\cos\theta_4 \\ -\cos\theta_5 & \cos(\theta_5+\theta_1) & \cos(\theta_3+\theta_4) & -\cos\theta_4 & 1 \end{pmatrix} \tag{3.5}$$

with $\boldsymbol{\theta} \in \mathbb{R}^5_+$ and $\mathbf{e}^\mathsf{T}\boldsymbol{\theta} < \pi$. Note that the Horn matrix H from (1.10) is of this form with $\boldsymbol{\theta} = \mathbf{0}$. It is easy to verify that H satisfies Property 3.4. The Horn matrix has been shown to be extremal for the cone of copositive matrices $\mathcal{COP}^5$, see [HN63].

The matrices $S(\boldsymbol{\theta})$ are transformed versions of the matrices $T(\boldsymbol{\varphi})$ that were introduced by Hildebrand in [Hil12] by making a trivial substitution. Observe, however, that we study $S(\boldsymbol{\theta})$ with $\boldsymbol{\theta}$ in a different range: While Hildebrand [Hil12] showed that an $S(\boldsymbol{\theta})$ with $\boldsymbol{\theta} \in \mathbb{R}^5_{++}$ and $\boldsymbol{e}^\mathsf{T}\boldsymbol{\theta} < \pi$ is extremal for $\mathcal{COP}^5$, we specifically consider in this section $S(\boldsymbol{\theta})$ with $\boldsymbol{\theta} \in \mathbb{R}^5_+ \setminus (\mathbb{R}^5_{++} \cup \{\mathbf{0}\})$ and $\boldsymbol{e}^\mathsf{T}\boldsymbol{\theta} < \pi$. We show that these matrices provide further counterexamples to Baumert's theorem. Moreover these matrices will also be of use later in this chapter.

Theorem 3.11. *Let $\boldsymbol{\theta} \in \mathbb{R}^5_+ \setminus (\mathbb{R}^5_{++} \cup \{\mathbf{0}\})$ such that $\mathbf{e}^\mathsf{T}\boldsymbol{\theta} < \pi$. Then the matrix $S(\boldsymbol{\theta})$ is in $\mathcal{COP}^5$, not extremal, not positive semidefinite, irreducible with respect to $\mathcal{N}^5$ and satisfies Property 3.10, but not Property 3.4.*

Proof. Without loss of generality we can always cycle the indices so we can assume that $\theta_1 = 0$. It is then easy to verify that

$$S(\boldsymbol{\theta}) = \mathbf{a}\mathbf{a}^\mathsf{T} + \operatorname{Diag}(\mathbf{d}) H \operatorname{Diag}(\mathbf{d})$$

where H is the Horn matrix (1.10),

$$\mathbf{a} = \begin{pmatrix} -\sin(\frac{1}{2}(\ \ \theta_2+\theta_3+\theta_4+\theta_5)) \\ \sin(\frac{1}{2}(\ \ \theta_2+\theta_3+\theta_4+\theta_5)) \\ -\sin(\frac{1}{2}(-\theta_2+\theta_3+\theta_4+\theta_5)) \\ \sin(\frac{1}{2}(-\theta_2-\theta_3+\theta_4+\theta_5)) \\ -\sin(\frac{1}{2}(-\theta_2-\theta_3-\theta_4+\theta_5)) \end{pmatrix}, \quad \text{and} \quad \mathbf{d} = \begin{pmatrix} \cos(\frac{1}{2}(\ \ \theta_2+\theta_3+\theta_4+\theta_5)) \\ \cos(\frac{1}{2}(\ \ \theta_2+\theta_3+\theta_4+\theta_5)) \\ \cos(\frac{1}{2}(-\theta_2+\theta_3+\theta_4+\theta_5)) \\ \cos(\frac{1}{2}(-\theta_2-\theta_3+\theta_4+\theta_5)) \\ \cos(\frac{1}{2}(-\theta_2-\theta_3-\theta_4+\theta_5)) \end{pmatrix}.$$

We have $\mathbf{d} > 0$ from which we see that $S(\boldsymbol{\theta})$ is copositive. We also have that $\mathbf{a} \neq \mathbf{0}$ from which we see that $S(\boldsymbol{\theta})$ is not extremal. For the set of zeros of $S(\boldsymbol{\theta})$ it can be seen that

$$\begin{aligned} \mathcal{V}^{S(\boldsymbol{\theta})} &= \operatorname{Diag}(\mathbf{d})^{-1} \mathcal{V}^{\operatorname{Diag}(\mathbf{d})^{-1} S(\boldsymbol{\theta}) \operatorname{Diag}(\mathbf{d})^{-1}} \\ &= \operatorname{Diag}(\mathbf{d})^{-1} \left(\mathcal{V}^{(\operatorname{Diag}(\mathbf{d})^{-1}\mathbf{a})(\operatorname{Diag}(\mathbf{d})^{-1}\mathbf{a})^\mathsf{T}} \cap \mathcal{V}^H \right) \\ &= \operatorname{Diag}(\mathbf{d})^{-1} \left\{ \mathbf{x} \in \mathcal{V}^H \;\middle|\; \mathbf{x}^\mathsf{T}(\operatorname{Diag}(\mathbf{d})^{-1}\mathbf{a}) = 0 \right\}, \end{aligned}$$

where we observe that

$$\mathcal{V}^H = \bigcup_{\substack{\text{cyclic} \\ \text{permutations}}} \operatorname{cone} \left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} \right\} \setminus \{\mathbf{0}\},$$

and

$$\operatorname{Diag}(\mathbf{d})^{-1}\mathbf{a} = \begin{pmatrix} -\tan(\frac{1}{2}(\ \ \theta_2+\theta_3+\theta_4+\theta_5)) \\ \tan(\frac{1}{2}(\ \ \theta_2+\theta_3+\theta_4+\theta_5)) \\ -\tan(\frac{1}{2}(-\theta_2+\theta_3+\theta_4+\theta_5)) \\ \tan(\frac{1}{2}(-\theta_2-\theta_3+\theta_4+\theta_5)) \\ -\tan(\frac{1}{2}(-\theta_2-\theta_3-\theta_4+\theta_5)) \end{pmatrix}.$$

If $\mathbf{u} \in \mathcal{V}^{S(\boldsymbol{\theta})}$ then we have that $S(\boldsymbol{\theta})\mathbf{u} = \operatorname{Diag}(\mathbf{d}) H \operatorname{Diag}(\mathbf{d})\mathbf{u}$. From this we observe that:

- If $\mathbf{u} \in \mathcal{V}^{S(\boldsymbol{\theta})}$ such that $\operatorname{supp}(\mathbf{u}) = \{1, 2\}$, then $\operatorname{supp}(S(\boldsymbol{\theta})\mathbf{u}) = \{4\}$.
- If $\mathbf{u} \in \mathcal{V}^{S(\boldsymbol{\theta})}$ such that $\operatorname{supp}(\mathbf{u}) = \{1, 2, 3\}$, then $\operatorname{supp}(S(\boldsymbol{\theta})\mathbf{u}) = \{4, 5\}$.

These results apply similarly after cyclic permutations.

For $i, j \in \{1, \ldots, 5\}$ we now say that (i, j) has Property I or Property II if and only if the following respective conditions hold:

I. There exists $\mathbf{u} \in \mathcal{V}^{S(\boldsymbol{\theta})}$ such that $u_i u_j > 0$.

II. There exists $\mathbf{u} \in \mathcal{V}^{S(\boldsymbol{\theta})}$ such that $u_i > 0 = u_j$ and $(S(\boldsymbol{\theta})\mathbf{u})_j = 0$.

We then have that $S(\boldsymbol{\theta})$ satifies Property 3.4 if and only if for all $i, j = 1, \ldots, 5$ we have that (i, j) satisfies Property I, and $S(\boldsymbol{\theta})$ satisfies Property 3.10 if and only if for all $i, j = 1, \ldots, 5$, $i < j$ at least one of (i, j) or (j, i) satisfies at least one of properties I or II.

Without loss of generality (by considering cyclic permutations) we now have 6 cases to consider. For each case, by looking at the structure of $\mathrm{Diag}(\mathbf{d})^{-1}\mathbf{a}$ and $\mathcal{V}^H$ it is a trivial but somewhat tedious task to find the support of the set of zeros. From this we can use the results above to check for each (i, j) if Property I or II holds. In each case we will name the case, give the support of the set of zeros and give a table showing for each (i, j) if Property I or II holds.

1. $\theta_1 = \theta_2 = \theta_3 = \theta_4 = 0$ and $0 < \theta_5 < \pi$: We have

$$\mathrm{supp}\left(\mathcal{V}^{S(\boldsymbol{\theta})}\right) = \Big\{\{1,2\},\{2,3\},\{3,4\},\{4,5\},\{1,2,3\},\{2,3,4\},\{3,4,5\}\Big\}$$

and

$j \backslash i$	1	2	3	4	5
1	I	I	I,II		II
2	I,II	I	I,II	I,II	II
3	I,II	I,II	I	I,II	I,II
4	II	I,II	I,II	I	I,II
5	II		I,II	I	I

2. $\theta_1 = \theta_2 = \theta_3 = 0$ and $0 < \theta_4, \theta_5$ and $\theta_4 + \theta_5 < \pi$: We have

$$\mathrm{supp}\left(\mathcal{V}^{S(\boldsymbol{\theta})}\right) = \Big\{\{1,2\},\{2,3\},\{3,4\},\{1,2,3\},\{2,3,4\},\{1,4,5\}\Big\}$$

and

$j \backslash i$	1	2	3	4	5
1	I	I	I,II	I	I,II
2	I,II	I	I,II	I,II	II
3	I,II	I,II	I	I,II	II
4	I	I,II	I	I	I,II
5	I			I	I

3. $\theta_1 = \theta_2 = \theta_4 = 0$ and $0 < \theta_3, \theta_5$ and $\theta_3 + \theta_5 < \pi$: We have

$$\operatorname{supp}\left(\mathcal{V}^{S(\boldsymbol{\theta})}\right) = \Big\{\{1,2\},\{2,3\},\{4,5\},\{1,2,3\}\Big\}$$

and

$j\backslash i$	1	2	3	4	5
1	I	I	I,II		II
2	I,II	I	I,II	II	II
3	I,II	I	I	II	
4	II		II	I	I
5	II		II	I	I

4. $\theta_1 = \theta_2 = 0$ and $0 < \theta_3, \theta_4, \theta_5$ and $\theta_3 + \theta_4 + \theta_5 < \pi$: We have

$$\operatorname{supp}\left(\mathcal{V}^{S(\boldsymbol{\theta})}\right) = \Big\{\{1,2\},\{2,3\},\{1,2,3\},\{2,3,4\},\{1,4,5\}\Big\}$$

and

$j\backslash i$	1	2	3	4	5
1	I	I	I,II	I	I,II
2	I,II	I	I,II	II	II
3	I,II	I	I	I,II	I
4	I		I	I	I
5	I		I	I	I

5. $\theta_1 = \theta_3 = 0$ and $0 < \theta_2, \theta_4, \theta_5$ and $\theta_2 + \theta_4 + \theta_5 < \pi$: We have

$$\operatorname{supp}\left(\mathcal{V}^{S(\boldsymbol{\theta})}\right) = \Big\{\{1,2\},\{3,4\},\{1,4,5\}\Big\}$$

and

$j\backslash i$	1	2	3	4	5
1	I	I	II	I	I,II
2	I	I	II		II
3		II	I	I	II
4	I	II	I	I	I,II
5	I			I	I

6. $\theta_1 = 0$ and $0 < \theta_2, \theta_3, \theta_4, \theta_5$ and $\theta_2 + \theta_3 + \theta_4 + \theta_5 < \pi$: We have

$$\operatorname{supp}\left(\mathcal{V}^{S(\boldsymbol{\theta})}\right) = \Big\{\{1,2\},\{2,3,4\},\{3,4,5\},\{1,4,5\}\Big\}$$

and

$j\backslash i$	1	2	3	4	5
1	I	I	II	I	I,II
2	I	I	I,II	I	II
3		I	I	I	I
4	I	I	I	I	I
5	I		I	I	I

From analyzing these results we see that in each case the matrix $S(\boldsymbol{\theta})$ satisfies Property 3.10 but not Property 3.4. Since (i,i) always has Property I it follows that $S(\boldsymbol{\theta})$ is irreducible with respect to $\mathcal{N}^5$.

Now let (i,j) have neither Property I nor Property II. From the tables above one can see that such a pair always exists. Let $\mathbf{u} \in \mathcal{V}^{S(\boldsymbol{\theta})}$ be such that $u_i > 0$. If now $S(\boldsymbol{\theta}) \in \mathcal{S}_+^5$, then $S(\boldsymbol{\theta})\mathbf{u} = \mathbf{0}$. Therefore, if $u_j > 0$, then (i,j) has Property I, and if $u_j = 0$, then (i,j) has Property II. Thus we obtain a contradiction, which proves that $S(\boldsymbol{\theta})$ is not positive semidefinite. This completes the proof. □

3.4 Auxiliary results

We next want to study the form of irreducible matrices. We start with two trivial, but important, lemmas.

Lemma 3.12. *For $n \geq 2$, let $A \in \mathcal{COP}^n$ and $i \in \{1,\dots,n\}$ such that $A_{ii} = 0$. Then A is irreducible with respect to $\widetilde{\mathcal{N}}^n$ if and only if $a_{ij} = 0$ for all $j = 1,\dots,n$, and $A_{\{1,\dots,n\}\setminus\{i\}}$ is irreducible with respect to $\widetilde{\mathcal{N}}^{n-1}$.*

Proof. This is trivial to see after recalling that every principle submatrix of a copositive matrix must be copositive. □

Lemma 3.13. *Let $A \in \mathcal{S}_+^n + \mathcal{N}^n$ be irreducible with respect to $\widetilde{\mathcal{N}}^n$. Then $A \in \mathcal{S}_+^n$.*

Proof. This is trivial to see from

$$\mathcal{N}^n \setminus \widetilde{\mathcal{N}}^n = \{\mathrm{Diag}(\mathbf{d}) \mid \mathbf{d} \in \mathbb{R}_+^n\} \setminus \{\mathbf{0}\} \subset \mathcal{S}_+^n$$

and $\mathcal{S}_+^n + \mathcal{N}^n \subseteq \mathcal{COP}^n$. □

Note again that for $n \leq 4$ we have $\mathcal{S}_+^n + \mathcal{N}^n = \mathcal{COP}^n$, so in these cases all of the matrices which are irreducible with respect to $\widetilde{\mathcal{N}}^n$ must be positive semidefinite. We will however now consider order 2 and order 3 copositive matrices more specifically. Due to Lemma 3.12 we may limit ourselves to cases when all the on-diagonal elements are strictly positive. By considering scaling with a diagonal matrix with on-diagonal elements strictly positive, we can in fact limit ourselves to the case where all the on-diagonal elements are equal to one.

Lemma 3.14. *Let $A \in \mathcal{COP}^2$ be such that* $\operatorname{diag}(A) = \mathbf{e}$*. Then we have that* $\mathcal{V}^A \subseteq \{\lambda \mathbf{e}_{12} \mid \lambda \in \mathbb{R}_{++}\}$ *and the following are equivalent:*

- *A is irreducible with respect to $\widetilde{\mathcal{N}}^2$,*
- $A = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$,
- $\mathcal{V}^A = \{\lambda \mathbf{e}_{12} \mid \lambda \in \mathbb{R}_{++}\}$.

Proof. For any $\mathbf{u} \in \mathbb{R}^2_+$ we have that $\mathbf{u}^\mathsf{T} A \mathbf{u} = u_1^2 + 2a_{12}u_1u_2 + u_2^2 = (u_1 - u_2)^2 + 2u_1u_2(a_{12}+1)$. Hence, for A to be copositive, we must have that $a_{12} \geq -1$. We also see that $\mathbf{u} \in \mathcal{V}^A$ if and only if $u_1 = u_2 > 0$ and $a_{12} = -1$. Now the fact that a matrix is irreducible with respect to $\widetilde{\mathcal{N}}^n$ if and only if Property 3.10 holds gives us the required results. ☐

The following corollary of this lemma is a consequence of the fact that all principal submatrices of a copositive matrix must be copositive, see property (ii) of Proposition 1.3.

Corollary 3.15. *Consider $A \in \mathcal{COP}^n$ with* $\operatorname{diag} A = \mathbf{e}$ *and let $\mathbf{u} \in \mathcal{V}^A$ be such that* $\operatorname{supp}(\mathbf{u}) = \{i, j\}$ *for some $i, j \in \{1, \dots, n\}$ with $i \neq j$. Then $a_{ij} = -1$ and there exists $\lambda \in \mathbb{R}_{++}$ such that $\mathbf{u} = \lambda \mathbf{e}_{ij}$.*

We now momentarily consider the following lemma.

Lemma 3.16. *Let $A \in \mathcal{COP}^n$ with* $\operatorname{diag} A = \mathbf{e}$ *be irreducible with respect to $\widetilde{\mathcal{N}}^n$. Then $a_{ij} \in [-1, 1]$ for all $i, j = 1, \dots, n$.*

Proof. Consider $A \in \mathcal{COP}^n$ such that $\operatorname{diag} A = \mathbf{e}$. All order 2 principle submatrices of A must be copositive, which implies that $a_{ij} \geq -1$ for all $i, j = 1, \dots, n$. Now considering work done on the copositive completion problem [HJR05] we see that all off-diagonal elements of A which are strictly greater than 1 can have their value replaced by 1 and the matrix would remain to be copositive. Therefore if A is irreducible with respect to $\widetilde{\mathcal{N}}^n$, then we must have that $a_{ij} \leq 1$ for all $i, j = 1, \dots, n$. ☐

Combining this lemma with Corollary 3.15 and noting that all order 3 principal submatrices of a copositive matrix must be in $\mathcal{COP}^3 = \mathcal{S}^3_+ + \mathcal{N}^3$ gives us the following result.

Lemma 3.17. *Let $A \in \mathcal{COP}^n$ be irreducible with respect to $\widetilde{\mathcal{N}}^n$ and* $\operatorname{diag} A = \mathbf{e}$*. Suppose $\{i,j\}, \{j,k\} \in \operatorname{supp}\left(\mathcal{V}^A\right)$, where i, j, k are mutually different. Then $A_{\{i,j,k\}}$ is a rank 1 positive semidefinite matrix with $1 = a_{ik} = -a_{ij} = -a_{jk}$.*

Next we consider order 3 copositive matrices specifically.

Lemma 3.18. *A matrix $A \in \mathcal{COP}^3$ with* $\operatorname{diag} A = \mathbf{e}$ *is irreducible with respect to $\widetilde{\mathcal{N}}^3$ if and only if it can be represented in the form*

$$A = \begin{pmatrix} 1 & -\cos\zeta & \cos(\zeta+\xi) \\ -\cos\zeta & 1 & -\cos\xi \\ \cos(\zeta+\xi) & -\cos\xi & 1 \end{pmatrix} \tag{3.6}$$

for some $(\zeta,\xi) \in \Psi := \{(\zeta,\xi) \in \mathbb{R}^2_+ \mid \zeta + \xi \leq \pi\}$.

We have that $\operatorname{supp}\left(\mathcal{V}^A\right) = \{\{1,2,3\}\}$ *if and only if $(\zeta,\xi) \in \operatorname{int}(\Psi)$. In this case the zero is proportional to $(\sin\xi, \sin(\zeta+\xi), \sin\zeta)^\mathsf{T}$.*

Proof. Before we begin this proof, first note that a matrix A of the form (3.6) can be decomposed as

$$A = \begin{pmatrix} -\cos\zeta \\ 1 \\ -\cos\xi \end{pmatrix} \begin{pmatrix} -\cos\zeta \\ 1 \\ -\cos\xi \end{pmatrix}^\mathsf{T} + \begin{pmatrix} \sin\zeta \\ 0 \\ -\sin\xi \end{pmatrix} \begin{pmatrix} \sin\zeta \\ 0 \\ -\sin\xi \end{pmatrix}^\mathsf{T} \in \mathcal{S}^3_+ \subset \mathcal{COP}^3. \tag{3.7}$$

Now let us show that if $A \in \mathcal{COP}^3$ with $\operatorname{diag} A = \boldsymbol{e}$ is irreducible with respect to $\widetilde{\mathcal{N}}^3$ then it must be represented in the required form. From Lemma 3.16 we see that there must exist $\zeta, \xi \in [0,\pi]$ and $a_{13} \in [-1,1]$ such that

$$A = \begin{pmatrix} 1 & -\cos\zeta & a_{13} \\ -\cos\zeta & 1 & -\cos\xi \\ a_{13} & -\cos\xi & 1 \end{pmatrix}.$$

As A is irreducible with respect to $\widetilde{\mathcal{N}}^3$, we must have that $\mathcal{V}^A \neq \emptyset$, and thus

$$0 = \det A = -(a_{13} - \cos\zeta\cos\xi)^2 + \sin^2\zeta\sin^2\xi.$$

Therefore $a_{13} = \cos\zeta\cos\xi \pm \sin\zeta\sin\xi$. For both possible values of a_{13} we would get that the matrix A is positive semidefinite, and thus copositive. Therefore A being irreducible with respect to $\widetilde{\mathcal{N}}^3$ means that

$$a_{13} = \min\{\cos\zeta\cos\xi \pm \sin\zeta\sin\xi\} = \cos\zeta\cos\xi - \sin\zeta\sin\xi = \cos(\zeta+\xi).$$

We are now left to show that $\zeta + \xi \leq \pi$. Suppose for the sake of contradiction that $\zeta + \xi > \pi$. We must have that $\zeta, \xi > 0$. Assuming $\zeta = \xi = \pi$ would give that A is the matrix with all entries equal to one, which is clearly not irreducible with respect to $\widetilde{\mathcal{N}}^3$. Therefore $\zeta + \xi < 2\pi$ and w.l.o.g. assume $\zeta \in (0, \pi)$. This implies that the vectors $(-\cos\zeta, 1, -\cos\xi)^\mathsf{T}$ and $(\sin\zeta, 0, -\sin\xi)^\mathsf{T}$ from (3.7) are linearly independent and so we get that

$$\begin{aligned}
\emptyset \neq \mathcal{V}^A &= \mathbb{R}^3_+ \cap \left\{ \mathbf{x} \in \mathbb{R}^3 \;\middle|\; 0 = \mathbf{x}^\mathsf{T} \begin{pmatrix} -\cos\zeta \\ 1 \\ -\cos\xi \end{pmatrix} = \mathbf{x}^\mathsf{T} \begin{pmatrix} \sin\zeta \\ 0 \\ -\sin\xi \end{pmatrix} \right\} \setminus \{\mathbf{0}\} \\
&= \mathbb{R}^3_+ \cap \left\{ \lambda \begin{pmatrix} \sin\xi \\ \sin(\zeta+\xi) \\ \sin\zeta \end{pmatrix} \;\middle|\; \lambda \in \mathbb{R} \setminus \{\mathbf{0}\} \right\} \\
&= \emptyset.
\end{aligned}$$

This gives us our contradiction, and so if $A \in \mathcal{COP}^3$ with $\operatorname{diag} A = \mathbf{e}$ is irreducible with respect to $\widetilde{\mathcal{N}}^3$ then it must be representable in the required form.

Next we show that any matrix of the form (3.6) is irreducible with respect to $\widetilde{\mathcal{N}}^3$. From the decomposition (3.7) we see that it is positive semidefinite, and hence copositive. If $\zeta = 0$ then $A\mathbf{e}_{12} = \mathbf{0}$ and so A is irreducible with respect to $\widetilde{\mathcal{N}}^3$ by Theorem 3.8. Similar reasoning holds for $\xi = 0$ and for $\zeta + \xi = \pi$, so we are left to consider $(\zeta, \xi) \in \operatorname{int}(\Psi)$. In this case we have that $\mathbf{u} := (\sin\xi, \sin(\zeta+\xi), \sin\zeta)^\mathsf{T} \in \mathcal{V}^A$ is a strictly positive vector, $A\mathbf{u} = \mathbf{0}$, and so again by Theorem 3.8 we have that A is irreducible with respect to $\widetilde{\mathcal{N}}^3$.

Finally from this discussion it is trivial to prove the last result for $\operatorname{supp}\left(\mathcal{V}^A\right) = \{\{1,2,3\}\}$. $\square$

From the last lemma we get the following corollary.

Corollary 3.19. *Let $A \in \mathcal{COP}^3$ with $\operatorname{diag} A = \mathbf{e}$ and $a_{13} = 1$ be irreducible with respect to $\widetilde{\mathcal{N}}^3$. Then $a_{12} = a_{23} = -1$ so that $\operatorname{supp}\left(\mathcal{V}^A\right) = \{\{1,2\},\{2,3\},\{1,2,3\}\}$.*

The main work of the following section will be on classifying matrices in $\mathcal{COP}^5$ which are irreducible with respect to $\widetilde{\mathcal{N}}^5$. Before we begin this, however, we will first need a few more auxilliary results.

Lemma 3.20 (Lemma 3.1 of [Bau66])**.** *Let $A \in \mathcal{COP}^n$ and $\boldsymbol{u} \in \mathcal{V}^A$ be such that $A\boldsymbol{u} = \mathbf{0}$. Let $\mathcal{I} = \operatorname{supp}(\mathbf{u})$. Then $A = P + C$, where $P \in \mathcal{S}^n_+$ is such that the rank of P equals the rank of the submatrix $P_\mathcal{I}$, and $C \in \mathcal{COP}^n$ is such that $c_{ij} = 0$ for all $i, j = 1, \ldots, n$ such that $j \in \mathcal{I}$.*

This important lemma has several consequences.

Corollary 3.21 (Corollary 3.2 of [Bau66])**.** *Let $A \in \mathcal{COP}^n$ and let $\mathbf{u} \in \mathcal{V}^A$ be such that $|\operatorname{supp}(\mathbf{u})| = n-1$ and $A\mathbf{u} = \mathbf{0}$. Then $A \in \mathcal{S}_+^n$.*

The following corollary comes trivially from combining Lemmas 3.13 and 3.20.

Corollary 3.22. *Let $A \in \mathcal{COP}^5$ be irreducible with respect to $\widetilde{\mathcal{N}}^5$ and $\mathbf{u} \in \mathcal{V}^A$ be such that $A\mathbf{u} = \mathbf{0}$. Then $A \in \mathcal{S}_+^5$.*

Next we consider the following lemma whose proof is similar to that of [Dia62, Lemma 11].

Lemma 3.23. *Let $A \in \mathcal{COP}^n$ be irreducible with respect to $\widetilde{\mathcal{N}}^n$, and assume there is an $(n-1)\times(n-1)$ principal submatrix B of A which is positive semidefinite. Then $A \in \mathcal{S}_+^n$.*

Proof. Assume w.l.o.g. that B is the upper left subblock and partition A as

$$A = \begin{pmatrix} B & \mathbf{a} \\ \mathbf{a}^\mathsf{T} & \mu \end{pmatrix} \qquad \text{with } B \in \mathcal{S}_+^{n-1}, \mathbf{a} \in \mathbb{R}^{n-1}, \mu \in \mathbb{R}_+.$$

Let $\gamma \geq 0$ be the largest number such that $\tilde{A} := A - \gamma E_{\{nn\}} \in \mathcal{COP}^n$ and set $\tilde{\mu} := \mu - \gamma$. Then $\tilde{A}$ is irreducible with respect to $E_{\{nn\}}$, and by [Bau66, Theorem 3.4] there exists $\tilde{\mathbf{u}} \in \mathbb{R}_+^{n-1}$ such that $\mathbf{u} = (\tilde{\mathbf{u}}^\mathsf{T}, 1)^\mathsf{T}$ is a zero of $\tilde{A}$. By Lemma 3.5 we then have

$$\tilde{A}\mathbf{u} = \begin{pmatrix} B\tilde{\mathbf{u}} + \mathbf{a} \\ \mathbf{a}^\mathsf{T}\tilde{\mathbf{u}} + \tilde{\mu} \end{pmatrix} \geq \mathbf{0}.$$

Moreover, by Lemma 3.7 we have that $(\tilde{A}\mathbf{u})_k = 0$ for all $k \in \operatorname{supp}(\mathbf{u})$. In particular, $0 = (\tilde{A}\mathbf{u})_n = \mathbf{a}^\mathsf{T}\tilde{\mathbf{u}} + \tilde{\mu}$ and $0 = \mathbf{u}^\mathsf{T}\tilde{A}\mathbf{u} - (\tilde{A}\mathbf{u})_n = \tilde{\mathbf{u}}^\mathsf{T}B\tilde{\mathbf{u}} + \mathbf{a}^\mathsf{T}\tilde{\mathbf{u}} = 0$. Therefore

$$A = \begin{pmatrix} B & -B\tilde{\mathbf{u}} \\ (-B\tilde{\mathbf{u}})^\mathsf{T} & \tilde{\mathbf{u}}^\mathsf{T}B\tilde{\mathbf{u}} \end{pmatrix} + \begin{pmatrix} 0 & B\tilde{\mathbf{u}} + \mathbf{a} \\ (B\tilde{\mathbf{u}} + \mathbf{a})^\mathsf{T} & \gamma \end{pmatrix}$$

which is the sum of a positive semidefinite matrix and a nonnegative matrix. The proof is concluded by Lemma 3.13. □

Combining this lemma with Lemma 3.13 and $\mathcal{S}_+^4 + \mathcal{N}^4 = \mathcal{COP}^4$ [Dia62, Theorem 2] gives us the following corollary.

Corollary 3.24. *Let $A \in \mathcal{COP}^5$ be irreducible with respect to $\widetilde{\mathcal{N}}^5$, and let some 4×4 submatrix B of A be irreducible with respect to $\widetilde{\mathcal{N}}^4$. Then $A \in \mathcal{S}_+^5$.*

Corollary 3.24 is an analogue of [Bau67, Lemma 4.4], where this assertion has been proven for irreducibility with respect to $\mathcal{N}^n$.

Corollary 3.25. *Let $A \in \mathcal{COP}^n$ be irreducible with respect to $\widetilde{\mathcal{N}}^n$. If there exists $\mathbf{u} \in \mathcal{V}^A$ such that $|\operatorname{supp}(\mathbf{u})| \geq n-1$, then $A \in \mathcal{S}_+^n$.*

Proof. If A has a zero with n positive components, then A is positive semidefinite by [Dia62, Lemma 7, (i)]. If A has a zero with $n-1$ positive components, then the corresponding principal submatrix is positive semidefinite by [Dia62, Lemma 7, (i)] and A is positive semidefinite by Lemma 3.23. $\square$

Corollary 3.26. *Let $A \in \mathcal{COP}^n$ be irreducible with respect to $\widetilde{\mathcal{N}}^n$. If for $i \neq j$, $E_{\{ij\}}$ is associated to a zero $\mathbf{u}$ with $|\operatorname{supp}(\mathbf{u})| \geq n-2$ and satisfying $u_i u_j = 0$, then $A \in \mathcal{S}_+^n$.*

Proof. Assume the condition of the corollary and let $\mathcal{I} = \operatorname{supp}(\mathbf{u})$. Define $\mathcal{I}' = \mathcal{I} \cup \{i,j\}$. Since $u_i u_j = 0$, but $u_i + u_j > 0$, the index set $\mathcal{I}'$ has at least $n-1$ elements. Moreover, by Definition 3.9 for association and by Lemma 3.7, we have $(A\mathbf{u})_k = 0$ for all $k \in \mathcal{I}'$, and therefore $A_{\mathcal{I}'}$ is positive semidefinite by Corollary 3.21. The proof is concluded by Lemma 3.23. $\square$

We finish this section with the following corollary.

Corollary 3.27. *Let $A \in \mathcal{COP}^5 \setminus \mathcal{S}_+^5$ with $\operatorname{diag} A = \mathbf{e}$ be irreducible with respect to $\widetilde{\mathcal{N}}^5$. Then every zero of A either has exactly 2 or exactly 3 positive components. If (i,j) is such that $u_i u_j = 0$ for all $\mathbf{u} \in \mathcal{V}^A$, then $E_{\{ij\}}$ is associated to a zero $\mathbf{e}_{kl}$ with $k \in \{i,j\}$ and $l \notin \{i,j\}$, and $0 = a_{ik} + a_{il} = a_{jk} + a_{jl}$.*

Proof. By Corollary 3.25, A cannot have a zero with more than 3 positive components. On the other hand, $\operatorname{diag} A = \mathbf{e}$ contradicts the existence of zeros with exactly one positive component.

Now if (i,j) is such that $u_i u_j = 0$ for all $\mathbf{u} \in \mathcal{V}^A$, then by Corollary 3.26, $E_{\{ij\}}$ cannot be associated to a zero with 3 positive components. Hence $E_{\{ij\}}$ is associated to a zero with precisely 2 positive components, which must have the support described in the assertion of the corollary and the condition $\operatorname{diag} A = \mathbf{e}$ ensures it is proportional to $\mathbf{e}_{kl}$, so it can be taken equal to. Finally by Definition 3.9 for association we have that $a_{ik} + a_{il} = a_{jk} + a_{jl} = 0$. $\square$

3.5 Classification of 5×5 copositive matrices

In this section we study matrices $A \in \mathcal{COP}^5$ which are irreducible with respect to $\widetilde{\mathcal{N}}^5$.

Let us first consider the case of A having a zero diagonal entry, say $a_{55} = 0$. Then for $k = 1, \ldots, 4$ we have by copositivity of A that $a_{k5} \geq 0$, and by irreducibility w.r.t. $E_{\{k5\}}$ we must have $a_{k5} = 0$. Hence A effectively is a copositive 4×4 matrix augmented with a zero row and column, which implies $A \in \mathcal{S}^5_+ + \mathcal{N}^5$ by [Dia62, Theorem 2]. By Lemma 3.13 we then have that A is positive semidefinite.

We may therefore assume that all diagonal elements of A are strictly positive. By possibly conjugating A with a positive definite diagonal matrix, we may assume without loss of generality that $\operatorname{diag} A = \mathbf{e}$. We next study irreducibility with respect to $\widetilde{\mathcal{N}}^5$, first when Property 3.4 does hold, and then when it does not.

3.5.1 Property 3.4

In [Bau67] the zero patterns of irreducible martices in $\mathcal{COP}^5$ which satisfy Property 3.4 have been classified. The result is summarized in the following lemma.

Lemma 3.28 (pp. 10–15 of [Bau67]). *Let $A \in \mathcal{COP}^5$ have* $\operatorname{diag} A = \mathbf{e}$ *and suppose A satisfies Property 3.4. Then either*

(a) A is positive semidefinite, or

(b) A is in the orbit of the Horn matrix H, or

(c) there exists a relabeling of variables such that

$$\{\{1,2,3\},\{1,2,4\},\{1,2,5\},\{3,4,5\}\} \subseteq \operatorname{supp}\left(\mathcal{V}^A\right),$$

or

(d) there exists a relabeling of variables such that

$$\{\{1,2,3\},\{2,3,4\},\{3,4,5\},\{1,4,5\},\{1,2,5\}\} = \operatorname{supp}\left(\mathcal{V}^A\right).$$

We shall now analyse this in detail and show that, in fact, case (c) is a subcase of (a), and case (d) means that A is in the orbit of some $S(\boldsymbol{\theta})$ from (3.5). We start with case (c).

Lemma 3.29. *Let $A \in \mathcal{COP}^5$ with* $\operatorname{diag} A = \mathbf{e}$ *be such that*

$$\{\{1,2,3\},\{1,2,4\},\{1,2,5\},\{3,4,5\}\} \subseteq \operatorname{supp}\left(\mathcal{V}^A\right)$$

Then A is positive semidefinite.

Proof. By Theorem 3.8 the matrix A and its submatrices $A_{\{1,2,3\}}, A_{\{1,2,4\}}$, $A_{\{1,2,5\}}$ and $A_{\{3,4,5\}}$ are irreducible with respect to $\widetilde{\mathcal{N}}^5$ resp. $\widetilde{\mathcal{N}}^3$ and by Lemma 3.6 these submatrices are positive semidefinite. By Lemma 3.18 we then have

$$A = \begin{pmatrix} 1 & -\cos\theta_1 & \cos(\theta_1+\theta_2) & \cos(\theta_1+\theta_3) & \cos(\theta_1+\theta_4) \\ -\cos\theta_1 & 1 & -\cos\theta_2 & -\cos\theta_3 & -\cos\theta_4 \\ \cos(\theta_1+\theta_2) & -\cos\theta_2 & 1 & -\cos\theta_5 & \cos(\theta_5+\theta_6) \\ \cos(\theta_1+\theta_3) & -\cos\theta_3 & -\cos\theta_5 & 1 & -\cos\theta_6 \\ \cos(\theta_1+\theta_4) & -\cos\theta_4 & \cos(\theta_5+\theta_6) & -\cos\theta_6 & 1 \end{pmatrix}$$

with

$$\boldsymbol{\theta} \in \mathbb{R}^6_+, \quad \theta_1+\theta_2 \le \pi, \quad \theta_1+\theta_3 \le \pi, \quad \theta_1+\theta_4 \le \pi, \quad \theta_5+\theta_6 \le \pi.$$

We shall now split the remainder of this proof in to three cases:

Case 1: $\theta_1 = 0$. It can be immediately seen that $\mathbf{e}_{12} \in \mathcal{V}^A$ and $A\mathbf{e}_{12} = \mathbf{0}$. Therefore Corollary 3.22 implies that A must be positive semidefinite.

Case 2: $\theta_1 = \pi$. From the inequalities on $\boldsymbol{\theta}$ we have that $0 = \theta_2 = \theta_3 = \theta_4$. We then get that $\mathbf{e}_{1k} \in \mathcal{V}^A$ for $k = 3,4,5$. From considering Lemma 3.5 we see that for $k,m = 3,4,5$ we have that $0 \le (A\boldsymbol{e}_{1k})_m = a_{1m} + a_{km} = -1 + a_{km}$. Therefore $a_{km} \ge 1$ for all $k,m = 3,4,5$, which by looking at the form of A gives us a contradiction and so this case is not possible.

Case 3: $0 < \theta_1 < \pi$. From the inequalities on $\boldsymbol{\theta}$ we have that $0 \le \theta_2, \theta_3, \theta_4 < \pi$ and $0 \le \theta_5, \theta_6, (\theta_5+\theta_6) \le \pi$.

If we let $\mathbf{u} := (\sin\theta_2, \sin(\theta_1+\theta_2), \sin\theta_1, 0, 0)^\mathsf{T}$ and $\mathbf{v} := (\sin\theta_3, \sin(\theta_1+\theta_3), 0, \sin\theta_1, 0)^\mathsf{T}$, then we have that $\mathbf{u}, \mathbf{v} \in \mathcal{V}^A$ and therefore by Lemma 3.5 $A\mathbf{u} \ge \mathbf{0}$ and $A\mathbf{v} \ge \mathbf{0}$. Specifically from using standard trigonometric identities we get that

$$0 \le (A\mathbf{u})_4 = -2\sin\theta_1 \cos\left(\tfrac{1}{2}(|\theta_2-\theta_3|+\theta_5)\right)\cos\left(\tfrac{1}{2}(|\theta_2-\theta_3|-\theta_5)\right).$$

We have $0 < \theta_1 < \pi$ and $0 \le |\theta_2-\theta_3| < \pi$ and $0 \le \theta_5 \le \pi$, and combining these with the inequalities above gives us that $|\theta_2-\theta_3| \ge \pi - \theta_5$. Similarly by considering $(A\mathbf{u})_5$ we get that $|\theta_2-\theta_4| \ge \theta_5+\theta_6$ and by considering $(A\boldsymbol{v})_5$ we see that $|\theta_3-\theta_4| \ge \pi-\theta_6$. Adding these new inequalities together we get

$$2\pi \le |\theta_2-\theta_3| + |\theta_2-\theta_4| + |\theta_3-\theta_4| = 2\max\{|\theta_2-\theta_3|, |\theta_2-\theta_4|, |\theta_3-\theta_4|\} < 2\pi,$$

a contradiction, so this case is not possible either. ☐

Next, we study case (d) of Lemma 3.28.

Lemma 3.30. *Let $A \in \mathcal{COP}^5$ with* $\operatorname{diag} A = \mathbf{e}$ *be such that*

$$\{\{1,2,3\},\{2,3,4\},\{3,4,5\},\{1,4,5\},\{1,2,5\}\} = \operatorname{supp}\left(\mathcal{V}^A\right).$$

Then $A = S(\boldsymbol{\theta})$ with $\boldsymbol{\theta} \in \mathbb{R}^5_{++}$ and $\mathbf{e}^\mathsf{T}\boldsymbol{\theta} < \pi$.

Proof. By considering Lemma 3.18 and Theorem 3.8 we can immediately see that $A = S(\boldsymbol{\theta})$ with $\boldsymbol{\theta} \in \mathbb{R}^5_{++}$ and $\theta_i + \theta_{(i+1) \bmod 5} < \pi$ for all $i = 1, \dots, 5$. We are now left to show that $\mathbf{e}^\mathsf{T}\boldsymbol{\theta} < \pi$.
First suppose that $\mathbf{e}^\mathsf{T}\boldsymbol{\theta} = \pi$. Then we have that

$$S(\boldsymbol{\theta}) = \begin{pmatrix} \cos(\theta_1+\theta_2) \\ -\cos\theta_2 \\ 1 \\ -\cos\theta_3 \\ \cos(\theta_3+\theta_4) \end{pmatrix} \begin{pmatrix} \cos(\theta_1+\theta_2) \\ -\cos\theta_2 \\ 1 \\ -\cos\theta_3 \\ \cos(\theta_3+\theta_4) \end{pmatrix}^\mathsf{T} + \begin{pmatrix} \sin(\theta_1+\theta_2) \\ -\sin\theta_2 \\ 1 \\ \sin\theta_3 \\ -\sin(\theta_3+\theta_4) \end{pmatrix} \begin{pmatrix} \sin(\theta_1+\theta_2) \\ -\sin\theta_2 \\ 1 \\ \sin\theta_3 \\ -\sin(\theta_3+\theta_4) \end{pmatrix}^\mathsf{T},$$

and from this we would get more than the required zeros, and thus $\boldsymbol{e}^\mathsf{T}\boldsymbol{\theta} = \pi$ is not possible.

Now suppose that $\mathbf{e}^\mathsf{T}\boldsymbol{\theta} > \pi$. We must have $\mathbf{e}^\mathsf{T}\boldsymbol{\theta} = \frac{1}{2}\sum_{i=1}^5(\theta_i + \theta_{(i+1) \bmod 5}) < \frac{5\pi}{2}$. As A is copositive, any order 4 principal submatrix of A must be copositive. We can in fact consider an arbitrary order 4 principal submatrix of A and then extend any results from this specific case to all order 4 principal submatrices of A by cycling the indices. So consider $\tilde{A} := A_{\{1,2,3,4\}}$ and $\mathbf{x} := (\sin\theta_2, \sin(\theta_1+\theta_2), \sin\theta_1, 0)^\mathsf{T}$. Then $\mathbf{x} \in \mathcal{V}^{\tilde{A}}$ and therefore by Lemma 3.5 we have that $\tilde{A}\mathbf{x} \geq \mathbf{0}$. In particular, using standard trigonometric identities, we get that

$$0 \leq (\tilde{A}\mathbf{x})_4 = 2\sin\theta_2\cos(\tfrac{1}{2}(\mathbf{e}^\mathsf{T}\boldsymbol{\theta}))\cos(\tfrac{1}{2}(\theta_1+\theta_2+\theta_3-\theta_4-\theta_5)).$$

We have that $0 < \theta_2 < \pi$ which implies that $\sin\theta_2 > 0$. We also have that $\frac{\pi}{2} < \frac{1}{2}(\mathbf{e}^\mathsf{T}\boldsymbol{\theta}) < \frac{5\pi}{4}$ which implies that $\cos(\frac{1}{2}(\mathbf{e}^\mathsf{T}\boldsymbol{\theta})) < 0$. This means that $\cos(\frac{1}{2}(\theta_1+\theta_2+\theta_3-\theta_4-\theta_5)) \leq 0$ which combined with the fact that $-\frac{\pi}{2} < \frac{1}{2}(\theta_1+\theta_2+\theta_3-\theta_4-\theta_5) < \frac{3\pi}{4}$ gives us that $\theta_1+\theta_2+\theta_3-\theta_4-\theta_5 \geq \pi$. We can now extend this result by cycling the indices to give us that

$$\theta_i + \theta_{(i+1) \bmod 5} + \theta_{(i+2) \bmod 5} - \theta_{(i+3) \bmod 5} - \theta_{(i+4) \bmod 5} \geq \pi \qquad \text{for all } i = 1, \dots, 5.$$

Adding these inequalities together then gives us the contradiction that $\mathbf{e}^\mathsf{T}\boldsymbol{\theta} \geq 5\pi$. ☐

We can now summarize the results of this subsection in to the following theorem.

Theorem 3.31. *Let $A \in \mathcal{COP}^5$ satisfy Property 3.4. Then either A is positive semidefinite, or A is in the orbit of $S(\boldsymbol{\theta})$ with $\boldsymbol{\theta} \in \mathbb{R}^5_{++} \cup \{\mathbf{0}\}$ and $\mathbf{e}^\mathsf{T}\boldsymbol{\theta} < \pi$.*

3.5.2 Property 3.10 but not Property 3.4

In this subsection we study when a matrix $A \in \mathcal{COP}^5$ is irreducible with respect to $\widetilde{\mathcal{N}}^5$ but does not satisfy Property 3.4, or in other words it satisfies Property 3.10 but not Property 3.4. The main result of this section will be to show that every such matrix must either be positive semidefinite or in the orbit of some $S(\boldsymbol{\theta})$ described in Theorem 3.11, i.e., a matrix $S(\boldsymbol{\theta})$ as in (3.5) with $\boldsymbol{\theta} \in \mathbb{R}^5_+ \setminus (\mathbb{R}^5_{++} \cup \{\mathbf{0}\})$ and $\mathbf{e}^\mathsf{T}\boldsymbol{\theta} < \pi$.

In order to prove this we shall assume, for the sake of contradiction, that there exists $A \in \mathcal{COP}^5$ such that A is irreducible with respect to $\widetilde{\mathcal{N}}^5$, Property 3.4 does not hold and A is neither positive semidefinite nor in the orbit of one of the matrices enumerated in Theorem 3.11. From the comment at the start of Section 3.5 we may again assume without loss of generality that $\operatorname{diag} A = \mathbf{e}$.

By Corollary 3.27, the zeros of A have exactly 2 or 3 positive components. Both types of zeros can be represented as edges in a graph with 5 vertices. We will represent a zero with support $\{i, j\}$ by a dashed edge between the vertices i and j, whilst we will represent a zero with support $\{1, \dots, 5\} \setminus \{i, j\}$ by a solid edge between the vertices i and j. We have that the graph of the zeros of A must fulfill the following conditions:

1. *For every vertex i there exist vertices j, k such that i, j, k are pairwise distinct, j, k are not linked by a dashed edge, and every solid edge having i as one of its vertices must have either j or k as its other vertex.* If for some i such vertices j, k do not exist, then the 4×4 submatrix obtained from A by crossing out row and column i satisfies Property 3.4. As a consequence, it satisfies also the weaker Property 3.10, and hence A is positive semidefinite by Theorem 3.8 and Corollary 3.24. In particular, *no three solid edges can join in a vertex, and no triangle consisting of two solid and one dashed edge is possible.*

2. *There exist distinct vertices i, j, not joined by a dashed edge, such that every solid edge has at least one of i, j as one of its vertices.* If such a pair i, j does not exist, then A satisfies Property 3.4.

3. *If two dashed edges join in a vertex, then the two vertices which do not intersect one of these dashed edges are joined by a solid edge.* By Lemma 3.17 the sum of the zeros represented by the dashed edges is also a zero, represented by the solid edge.

4. *If there is a dashed edge (i, j) and a solid edge (k, l) such that i, j, k, l are pairwise distinct, then there is another dashed edge joining either i, m or j, m, where m is the remaining vertex.* The solid edge stands for a $\mathbf{v} \in \mathcal{V}^A$ with support $\{i, j, m\}$, implying that the submatrix $A_{\{i,j,m\}}$ is positive semidefinite. But then $\mathbf{e}_{ij}$ and $\mathbf{v}$ are linearly independent kernel vectors of $A_{\{i,j,m\}}$, and

$A_{\{i,j,m\}}$ is of rank 1. Existence of either a zero $\mathbf{e}_{im}$ or $\mathbf{e}_{jm}$ now easily follows.

5. *There are no dashed triangles.* If there were a dashed triangle on the vertices i, j, k, then the submatrix $A_{\{i,j,k\}}$ would have all off-diagonal elements equal to -1 and could not be in $\mathcal{COP}^3$.

6. *There cannot exist pairwise distinct vertices i, j, k, l such that both (i, j) and (i, k) are dashed edges whilst (i, l) is a solid edge.* The existence of zeros $\mathbf{e}_{ij}, \mathbf{e}_{ik}$ implies by Lemma 3.17 that $a_{jk} = 1$. Since A has a zero with support $\{j, k, m\}$, where m is the remaining vertex, the submatrix $A_{\{j,k,m\}}$ is irreducible by Theorem 3.8. Existence of the zeros $\mathbf{e}_{jm}, \mathbf{e}_{km}$ now follows from Lemma 3.19. Now if we consider condition 3 we get (j, l), (k, l) and (l, m) are solid edges. Finally considering condition 1, we get a contradiction.

7. *For every two distinct vertices i, j, either there exists a dashed edge with at least one of i, j as one of its vertices, or there exists a solid edge whose vertices are not in $\{i, j\}$.* If a dashed edge with the specified properties does not exist, then $E_{\{ij\}}$ is not associated to a zero with exactly 2 positive components. Hence $E_{\{ij\}}$ must be associated to a zero $\mathbf{u}$ with 3 positive components, and by Corollary 3.27 this zero must satisfy $u_i u_j > 0$.

Applying these rules it can be found that the only graphs satisfying these conditions, up to permutation of the vertices, are given in Fig. 3.1. These give us 14 possible cases for A to consider, which we have ordered and permuted for ease of going through them. For each one we shall find a contradiction in the form of A being positive semidefinite or A being in the orbit of some $S(\boldsymbol{\theta})$ described in Theorem 3.11, i.e., $S(\boldsymbol{\theta})$ as in (3.5) with $\boldsymbol{\theta} \in \mathbb{R}^5_+ \setminus (\mathbb{R}^5_{++} \cup \{\mathbf{0}\})$ and $\mathbf{e}^\mathsf{T}\boldsymbol{\theta} < \pi$.

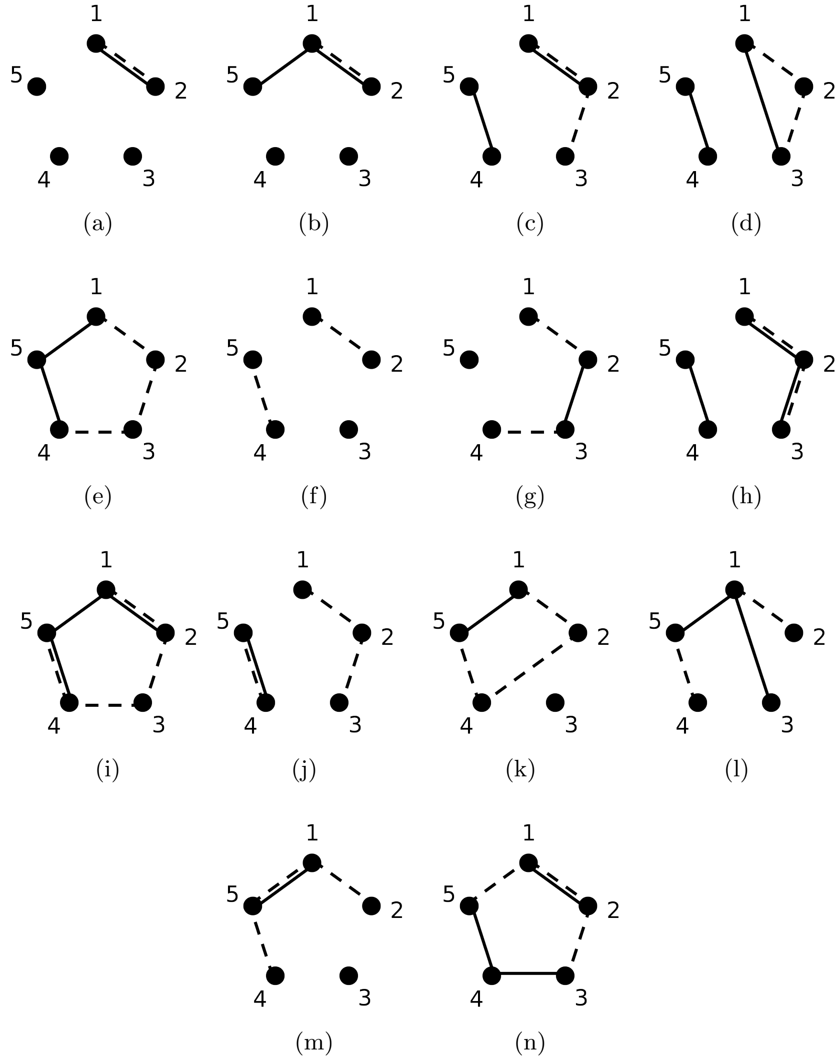

Figure 3.1: Graphs fulfilling the conditions in Subsection 3.5.2, used to represent the support of the set of zeros of a matrix. A dashed edge between vertices i and j represents a zero with support $\{i, j\}$, whilst a solid edge between the vertices i and j represents a zero with support $\{1, \dots, 5\} \setminus \{i, j\}$.

Before we do this however, we will first recall the following which we shall use regularly when going through these cases.

- By Corollary 3.15 we have that if $\{i, j\} \in \operatorname{supp}\left(\mathcal{V}^A\right)$ with $i \neq j$, then up to positive scalings $\mathbf{e}_{ij}$ is the unique zero with this support.

- By Lemma 3.7, for $\mathbf{u}, \mathbf{v} \in \mathcal{V}^A$ with $\mathrm{supp}(\mathbf{u}) \subseteq \mathrm{supp}(\mathbf{v})$ we have $(A\mathbf{u})_i = 0$ for all $i \in \mathrm{supp}(\mathbf{v})$.
- By Corollary 3.22, there cannot exist $\mathbf{u} \in \mathcal{V}^A$ such that $A\mathbf{u} = \mathbf{0}$.
- By Corollary 3.27, if $i \neq j$ is such that $u_i u_j = 0$ for all $\mathbf{u} \in \mathcal{V}^A$, then $E_{\{ij\}}$ is associated to a zero $\mathbf{e}_{kl}$ with $k \in \{i, j\}$ and $l \notin \{i, j\}$.
- By Lemma 3.16 we have that $a_{ij} \in [-1, 1]$ for all $i = 1, \ldots, 5$.

We shall use these results regularly whilst going through the cases without specifically referencing them.

(a) $\mathrm{supp}\left(\mathcal{V}^A\right) = \{\{3,4,5\}, \{1,2\}\}$:

We have $(A\mathbf{e}_{12})_i = 0$ for $i = 1, 2$. For $k = 3, 4, 5$ we must have that $E_{\{1k\}}$ is associated to $\mathbf{e}_{12}$, and thus $\mathbf{0} = A\mathbf{e}_{12} \neq \mathbf{0}$.

(b) $\mathrm{supp}\left(\mathcal{V}^A\right) = \{\{2,3,4\}, \{3,4,5\}, \{1,2\}\}$:

By following the same steps as in the previous case, we again get a contradiction.

(c) $\mathrm{supp}\left(\mathcal{V}^A\right) = \{\{1,2\}, \{2,3\}, \{1,2,3\}, \{3,4,5\}\}$:

We have $(A\mathbf{e}_{12})_i = 0$ for $i = 1, 2, 3$. We also have that both $E_{\{14\}}$ and $E_{\{15\}}$ must be associated to $\mathbf{e}_{12}$, hence $\mathbf{0} = A\mathbf{e}_{12} \neq \mathbf{0}$.

(d) $\mathrm{supp}\left(\mathcal{V}^A\right) = \{\{1,2\}, \{2,3\}, \{1,2,3\}, \{2,4,5\}\}$:

By following the same steps as in the previous case, we again get a contradiction.

(e) $\mathrm{supp}\left(\mathcal{V}^A\right) = \{\{1,2\}, \{2,3\}, \{3,4\}, \{1,2,3\}, \{2,3,4\}\}$:

We must have that $E_{\{14\}}$ is associated to either $\mathbf{e}_{12}$ or $\mathbf{e}_{34}$. Without loss of generality, let it be associated to $\mathbf{e}_{12}$, and so $(A\mathbf{e}_{12})_4 = 0$. However, we must also have $0 = (A\mathbf{e}_{12})_i$ for $i = 1, 2, 3$. Finally we have that $E_{\{15\}}$ must be associated to $\mathbf{e}_{12}$, implying $\mathbf{0} = A\mathbf{u}_{12} \neq \mathbf{0}$.

(f) $\mathrm{supp}\left(\mathcal{V}^A\right) = \{\{1,2\}, \{4,5\}\}$:

Without loss of generality $E_{\{24\}}$ is associated to $\mathbf{e}_{12}$, implying $(A\mathbf{e}_{12})_4 = 0$. We must have that $E_{\{13\}}$ and $E_{\{34\}}$ are associated to $\mathbf{e}_{12}$ and $\mathbf{e}_{45}$ respectively, hence $0 = (A\mathbf{e}_{12})_3 = (A\mathbf{e}_{45})_3$. We must then have $(A\mathbf{e}_{12})_5 > 0$, otherwise $A\mathbf{e}_{12} = \mathbf{0}$. This however implies that both $E_{\{15\}}$ and $E_{\{25\}}$ must be associated to $\mathbf{e}_{45}$ and so $\mathbf{0} = A\mathbf{e}_{45} \neq \mathbf{0}$.

(g) $\operatorname{supp}(\mathcal{V}^A) = \{\{1,2\},\{3,4\},\{1,4,5\}\}$:

Without loss of generality $E_{\{23\}}$ is associated to $\mathbf{e}_{12}$, implying $0 = (A\mathbf{e}_{12})_3 = a_{13} + a_{23}$. Further, $E_{\{25\}}$ and $E_{\{35\}}$ are associated to $\boldsymbol{e}_{12}$ and $\mathbf{e}_{34}$ respectively, implying that $0 = (A\mathbf{e}_{12})_5 = a_{15} + a_{25}$ and $0 = (A\mathbf{e}_{34})_5 = a_{35}+a_{45}$. We must have $(A\mathbf{e}_{12})_4 > 0$, otherwise $A\mathbf{e}_{12} = \mathbf{0}$, and hence $E_{\{24\}}$ must be associated to $\mathbf{e}_{34}$, implying $0 = (A\mathbf{e}_{34})_2 = a_{23}+a_{24}$. From this and Lemma 3.18 we see that

$$A = \begin{pmatrix} 1 & -1 & a_{13} & \cos(\theta_4+\theta_5) & -\cos\theta_5 \\ -1 & 1 & -a_{13} & a_{13} & \cos\theta_5 \\ a_{13} & -a_{13} & 1 & -1 & \cos\theta_4 \\ \cos(\theta_4+\theta_5) & a_{13} & -1 & 1 & -\cos\theta_4 \\ -\cos\theta_5 & \cos\theta_5 & \cos\theta_4 & -\cos\theta_4 & 1 \end{pmatrix} = S(\boldsymbol{\theta})$$

for some $\boldsymbol{\theta} = (0,\theta_2,0,\theta_4,\theta_5)$, where $0 < \theta_2,\theta_4,\theta_5,(\theta_4+\theta_5) < \pi$. We now recall the inequality $0 < (A\boldsymbol{e}_{12})_4 = \cos(\theta_4+\theta_5) + \cos\theta_2 = 2\cos(\frac{1}{2}(\theta_2 + \theta_4+\theta_5))\cos(\frac{1}{2}(\theta_4+\theta_5-\theta_2)$, which implies $\pi > \theta_2+\theta_4+\theta_5 = \mathbf{e}^\mathsf{T}\boldsymbol{\theta}$, and thus A is of the form $S(\boldsymbol{\theta})$ as in Theorem 3.11.

(h) $\operatorname{supp}(\mathcal{V}^A) = \{\{1,2\},\{2,3\},\{1,2,3\},\{3,4,5\},\{1,4,5\}\}$:

Without loss of generality we have that $E_{\{25\}}$ is associated to $\mathbf{e}_{12}$, implying $0 = (A\mathbf{e}_{12})_5 = a_{15} + a_{25}$. Note that $(A\mathbf{e}_{12})_i = 0$ for $i = 1,2,3$. Therefore $(A\mathbf{e}_{12})_4 > 0$, and hence $E_{\{24\}}$ must be associated to $\mathbf{e}_{23}$, implying $0 = (A\mathbf{e}_{23})_4 = a_{24}+a_{34}$. From this and Lemma 3.18 we now see that $A = S(\boldsymbol{\theta})$ for some $\boldsymbol{\theta} = (0,0,\theta_3,\theta_4,\theta_5)^\mathsf{T}$ with $0 < \theta_3,\theta_4,\theta_5,(\theta_3+\theta_4),(\theta_4+\theta_5) < \pi$. We recall $0 < (A\mathbf{e}_{12})_4 = 2\cos(\frac{1}{2}(\theta_3+\theta_4-\theta_5))\cos(\frac{1}{2}(\theta_3+\theta_4+\theta_5))$ from which we get that $\pi > \theta_3+\theta_4+\theta_5 = \mathbf{e}^\mathsf{T}\boldsymbol{\theta}$, and thus A is of the form $S(\boldsymbol{\theta})$ as in Theorem 3.11.

(i) $\operatorname{supp}(\mathcal{V}^A) = \{\{1,2\},\{2,3\},\{3,4\},\{4,5\},\{1,2,3\},\{2,3,4\},\{3,4,5\}\}$:

We have that

$$A = \begin{pmatrix} 1 & -1 & 1 & a_{14} & a_{15} \\ -1 & 1 & -1 & 1 & a_{25} \\ 1 & -1 & 1 & -1 & 1 \\ a_{14} & 1 & -1 & 1 & -1 \\ a_{15} & a_{25} & 1 & -1 & 1 \end{pmatrix} = S(\boldsymbol{\theta}) + \delta_1 E_{\{14\}} + \delta_2 E_{\{25\}}$$

for some $\boldsymbol{\theta} = (0,0,0,0,\theta_5)^\mathsf{T}$ and $\boldsymbol{\delta} \in \mathbb{R}^2$, with $0 < \theta_5 \le \pi$. Moreover, $0 \le (A\mathbf{e}_{45})_1 = \delta_1$ and $0 \le (A\mathbf{e}_{12})_5 = \delta_2$. Therefore A can be written as the sum of the nonnegative matrix $(\delta_1 E_{\{45\}}+\delta_2 E_{\{25\}})$ and the copositive matrix $S(\boldsymbol{\theta})$ from (3.5). As A is irreducible we must therefore have that

$\boldsymbol{\delta} = \mathbf{0}$, and thus A is either positive semidefinite or of the form $S(\boldsymbol{\theta})$ as in Theorem 3.11.

(j) $\operatorname{supp}\left(\mathcal{V}^A\right) = \{\{1,2\},\{2,3\},\{4,5\},\{1,2,3\}\}$:

We have that $0 = (A\mathbf{e}_{12})_i = (A\mathbf{e}_{23})_i$ for $i = 1,2,3$. If both $E_{\{14\}}$ and $E_{\{15\}}$ are associated to $\mathbf{e}_{12}$ then we would have that $A\mathbf{e}_{12} = \mathbf{0}$. Therefore, without loss of generality $0 < (A\mathbf{e}_{12})_4 = a_{14} + a_{24}$ and hence $E_{\{14\}}$ is associated to $\mathbf{e}_{45}$, implying that $0 = (A\mathbf{e}_{45})_1 = a_{14} + a_{15}$. Similarly, we cannot have $0 = (A\mathbf{e}_{23})_4 = (A\mathbf{e}_{23})_5$, and hence there must be $k \in \{4,5\}$ such that $(A\mathbf{e}_{23})_k > 0$. But then $E_{\{3k\}}$ must be associated to $\mathbf{e}_{45}$, and $0 = (A\mathbf{e}_{45})_3 = a_{34} + a_{35}$. $A\mathbf{e}_{45} \neq \mathbf{0}$ then yields $0 < (A\mathbf{e}_{45})_2$ and hence $E_{\{24\}}$ must be associated to $\mathbf{e}_{23}$, implying $0 = (A\mathbf{e}_{23})_4 = a_{24} + a_{34}$. $A\mathbf{e}_{23} \neq \mathbf{0}$ then yields $0 < (A\mathbf{e}_{23})_5$, and hence $E_{\{25\}}$ must be associated to $\mathbf{e}_{12}$, implying $0 = (A\mathbf{e}_{12})_5 = a_{15} + a_{25}$. Therefore

$$A = \begin{pmatrix} 1 & -1 & 1 & -a_{15} & a_{15} \\ -1 & 1 & -1 & -a_{34} & -a_{15} \\ 1 & -1 & 1 & a_{34} & -a_{34} \\ -a_{15} & -a_{34} & a_{34} & 1 & -1 \\ a_{15} & -a_{15} & -a_{34} & -1 & 1 \end{pmatrix} = S(\boldsymbol{\theta}),$$

for some $\boldsymbol{\theta} = (0,0,\theta_3,0,\theta_5)$ with $0 < \theta_3, \theta_5 \leq \pi$. We now recall the inequality $0 < a_{14} + a_{24} = 2\cos(\frac{1}{2}(\theta_3 - \theta_5))\cos(\frac{1}{2}(\theta_3 + \theta_5))$ which implies that $\pi > \theta_3 + \theta_5 = \mathbf{e}^\mathsf{T}\boldsymbol{\theta}$, and thus A is of the form $S(\boldsymbol{\theta})$ as in Theorem 3.11.

(k), (l), (m), (n) $\operatorname{supp}\left(\mathcal{V}^A\right) \supseteq \{\{1,5\},\{1,2,3\},\{3,4,5\}\}$:

By Lemma 3.18 we see that $A = S(\boldsymbol{\theta}) + \delta_1 E_{\{14\}} + \delta_2 E_{\{25\}} + \delta_3 E_{\{24\}}$, for some $\boldsymbol{\theta} = (\theta_1, \theta_2, \theta_3, \theta_4, 0)^\mathsf{T}$ and $\boldsymbol{\delta} \in \mathbb{R}^3$ such that $0 \leq \theta_1, \theta_2, \theta_3, \theta_4, (\theta_1 + \theta_2), (\theta_3 + \theta_4) \leq \pi$. By Lemma 3.5 we have $0 \leq (A\mathbf{e}_{15})_4 = \delta_1$ and $0 \leq (A\mathbf{e}_{15})_2 = \delta_2$. Moreover we have that $0 \leq (A\mathbf{e}_{15})_3 = 2\cos(\frac{1}{2}(\theta_1+\theta_2-\theta_3-\theta_4))\cos(\frac{1}{2}(\theta_1+\theta_2+\theta_3+\theta_4))$, implying that $\pi \geq \theta_1+\theta_2+\theta_3+\theta_4 = \mathbf{e}^\mathsf{T}\boldsymbol{\theta}$. Also considering the copositivity of

$$A_{\{2,3,4\}} = \begin{pmatrix} 1 & -\cos\theta_2 & \cos(\theta_2+\theta_3)+\delta_3 \\ -\cos\theta_2 & 1 & -\cos\theta_3 \\ \cos(\theta_2+\theta_3)+\delta_3 & -\cos\theta_3 & 1 \end{pmatrix}$$

along with Lemma 3.18 and the constraint $\theta_2 + \theta_3 \leq \pi$ yields $\delta_3 \geq 0$. Therefore A can be written as the sum of the nonnegative matrix $(\delta_1 E_{\{14\}} + \delta_2 E_{\{25\}} + \delta_3 E_{\{24\}})$ and a copositive matrix in the orbit of some $S(\boldsymbol{\theta})$ from (3.5). As A is irreducible we must therefore have that $\boldsymbol{\delta} = \mathbf{0}$, and thus A is of the form $S(\boldsymbol{\theta})$ as in Theorem 3.11.

Summing up, we have proven the following result.

Theorem 3.32. *Let $A \in \mathcal{COP}^5$ be irreducible with respect to $\widetilde{\mathcal{N}}^5$ and assume A does not satisfy Property 3.4. Then either A is positive semidefinite, or A is in the orbit of $S(\boldsymbol{\theta})$ with $\boldsymbol{\theta} \in \mathbb{R}^5_+ \setminus (\mathbb{R}^5_{++} \cup \{$
$mathbf 0\})$ and $\mathbf{e}^\mathsf{T}\boldsymbol{\theta} < \pi$.*

3.5.3 Irreducible matrices of $\mathcal{COP}^5$

By combining Theorems 3.31 and 3.32, we obtain the following result.

Theorem 3.33. *Let $A \in \mathcal{COP}^5$ be irreducible with respect to $\widetilde{\mathcal{N}}^5$. Then either A is positive semidefinite, or A is in the orbit of $S(\boldsymbol{\theta})$ given by (3.5) with $\boldsymbol{\theta} \in \mathbb{R}^5_+$ and $\boldsymbol{e}^\mathsf{T}\boldsymbol{\theta} < \pi$.*

As being irreducible with respect to $\mathcal{N}^5$ is a stronger statement, this theorem must also hold for irreducibility with respect to $\mathcal{N}^5$. The following theorem is also clear to see.

Theorem 3.34. *Let $A \in \mathcal{COP}^5$ be irreducible with respect to $\widetilde{\mathcal{N}}^5$, but not irreducible with respect to $\mathcal{N}^5$. Then A is positive semidefinite.*

Proof. Suppose for the sake of contradiction that there exists a matrix $A \in \mathcal{COP}^5 \setminus \mathcal{S}^5_+$ such that A is irreducible with respect to $\widetilde{\mathcal{N}}^5$, but not irreducible with respect to $\mathcal{N}^5$. It can be seen that A cannot satisfy Property 3.4, otherwise it would be irreducible with respect to $\mathcal{N}^5$. Therefore, by Theorem 3.32 and $A \notin \mathcal{S}^5_+$, we get that A must be in the orbit of $S(\boldsymbol{\theta})$ with $\boldsymbol{\theta} \in \mathbb{R}^5_+ \setminus (\mathbb{R}^5_{++} \cup \{\mathbf{0}\})$ and $\boldsymbol{e}^\mathsf{T}\boldsymbol{\theta} < \pi$. However, by Theorem 3.11, this means that A is irreducible with respect to $\mathcal{N}^5$, giving us a contradiction. $\square$

Our results immediately yield a simple characterization of those 5×5 copositive matrices which cannot be written as a sum of a positive semidefinite and a nonnegative matrix. Namely, we have the following result which, as mentioned before, can be seen as the dual statement to [BAD09, Corollary 2].

Corollary 3.35. *Let $A \in \mathcal{COP}^5 \setminus (\mathcal{S}^5_+ + \mathcal{N}^5)$. Then A can be expressed as $A = S + N$ for some $N \in \widetilde{\mathcal{N}}^5$ and $S \in \mathcal{COP}^5$ in the orbit of $S(\boldsymbol{\theta})$ given by (3.5), where $\boldsymbol{\theta} \in \mathbb{R}^5_+$ and $\mathbf{e}^\mathsf{T}\boldsymbol{\theta} < \pi$.*

Chapter 4

Scaling relationship between the copositive cone and Parrilo's first level approximation [1]

[1]Published as [DDGH13b] Peter J.C. Dickinson, Mirjam Dür, Luuk Gijben and Roland Hildebrand. Scaling relationship between the copositive cone and Parrilo's first level approximation. *Optimization Letters*, 7(8):1669-1679, 2013.

For the purpose of this chapter recall from (1.14) that the hierarchy of inner approximations for the copositive cone introduced by Parrilo [Par00] is defined as

$$\mathcal{K}_n^r = \Big\{ A \in \mathbb{S}^n \mid \Big(\sum_{i=1}^n x_i^2 \Big)^r (\mathbf{x} \circ \mathbf{x})^\mathsf{T} A (\mathbf{x} \circ \mathbf{x}) \text{ is SOS } \Big\}.$$

For $\mathcal{K}_n^1$, the combined proofs of Parrilo [Par00] (sufficient) and Bomze and de Klerk [BdK02] (necessary) show that $A \in \mathcal{K}_n^1$ if and only if the following system of LMIs has a feasible solution $M^1, \dots, M^n$:

$$A - M^i \in \mathcal{S}_+^n \qquad i = 1, \dots, n \tag{4.1a}$$

$$(M^i)_{ii} = 0 \qquad i = 1, \dots, n \tag{4.1b}$$

$$(M^i)_{jj} + 2(M^j)_{ij} = 0 \qquad i, j = 1, \dots, n \quad \text{s.t. } i \neq j \tag{4.1c}$$

$$(M^i)_{jk} + (M^j)_{ik} + (M^k)_{ij} \geq 0 \quad i, j, k = 1, \dots, n \quad \text{s.t. } i < j < k. \tag{4.1d}$$

Note that these LMIs can be directly obtained from the alternative definition (1.18) of $\mathcal{K}_n^r$ that was given in [PVZ07].

Furthermore recall again that $\mathcal{K}_n^0 = \mathcal{S}_+^n + \mathcal{N}^n$ and that $\mathcal{COP}^n = \mathcal{K}_n^0$ if and only if $n \leq 4$. In the 5×5 case, as was noted before, the Horn-matrix (1.10) is in $\mathcal{COP}^5 \setminus \mathcal{K}_5^0$, moreover an explicit certificate was given by Parrilo [Par00] that shows that the Horn matrix in fact is in $\mathcal{K}_5^1$. The natural question that was then posed was to determine the smallest n for which $\mathcal{COP}^n \neq \mathcal{K}_n^1$. In this chapter, we answer this question and show that in fact already $\mathcal{COP}^5 \neq \mathcal{K}_5^r$ for all $r \in \mathbb{Z}_+$. A central ingredient of the proof that we will present below is the observation that if we are given a diagonal matrix D with strictly positive diagonal, then for any matrix class $\mathcal{X} \in \{\mathcal{COP}^n, \mathcal{S}_+^n, \mathcal{N}^n, \mathcal{S}_+^n + \mathcal{N}^n\}$ we have that $A \in \mathcal{X} \iff DAD \in \mathcal{X}$, see Propositions 1.2 and 1.3. However, as we will see later on this property does not hold for $\mathcal{X} = \mathcal{K}_n^r$ when $r \geq 1$. We will show in fact that for any matrix $A \in \mathcal{COP}^n \setminus (\mathcal{S}_+^n + \mathcal{N}^n)$ and for any $r \in \mathbb{Z}_+$ there exists a diagonal matrix D with strictly positive diagonal, such that $DAD \notin \mathcal{K}_n^r$. In general we can only show that such scaling matrices D exist, but we do not know how to construct them. For the case when $r = 1$

however we give an explicit way to construct a scaling matrix that scales a given matrix in $\mathcal{K}_n^1 \setminus (\mathcal{S}_+^n + \mathcal{N}^n)$ out of $\mathcal{K}_n^1$. Furthermore we provide some remarks concerning scalings in the opposite direction, that is scaling such that $DAD \in \mathcal{K}_n^r$ for a matrix $A \in \mathcal{COP}^n \setminus \mathcal{K}_n^r$.

For the 5×5 case, we will show the more surprising result that for any matrix $A \in \mathcal{COP}^5$ scaling it in such a way that $(DAD)_{ii} \in \{0, 1\}$ will yield $DAD \in \mathcal{K}_5^1$. Our main result of this chapter (Theorem 4.18) is a complete characterization of $\mathcal{COP}^5$ in terms of $\mathcal{K}_5^1$. We will conclude this chapter by formulating several conjectures and open problems regarding scalings of copositive matrices with respect to the hierarchy of Parrilo cones.

Notation

We recall that given a vector $\mathbf{d} \in \mathbb{R}^n$, we denote by $\mathrm{Diag}(\mathbf{d})$ the diagonal matrix with the entries of $\mathbf{d}$ on its diagonal. Conversely, given a matrix A, we denote by $\mathrm{diag}(A)$ the vector of diagonal entries of A. We shall denote the set of scalings by

$$\mathcal{D} := \{\mathrm{Diag}(\mathbf{d}) \mid \mathbf{d} \in \mathbb{R}_{++}^n\}.$$

During this chapter we will use the modulo operator in such a way that it maps to $\{1, \ldots, n\}$ rather than $\{0, \ldots, n-1\}$. That is, $n \bmod n \equiv 0 \bmod n \equiv n$, and $i \bmod n \equiv i$ for $0 < i < n$. The reason for this is that it will make notation much more convenient later on improving readability.

4.1 Scaling a matrix out of $\mathcal{K}_n^r$

In this section we will show that for $n \geq 5$ we have that $\mathcal{K}_n^r \neq \mathcal{COP}^n$ for all $r \geq 0$. Instead of giving a specific example of a matrix in $\mathcal{COP}^n$ but not in $\mathcal{K}_n^r$, we will show that in fact any matrix $A \in \mathcal{COP}^n \setminus (\mathcal{S}_+^n + \mathcal{N}^n)$ can be "scaled out" of $\mathcal{K}_n^r$. As $\mathcal{S}_+^n + \mathcal{N}^n \neq \mathcal{COP}^n$ if and only if $n \geq 5$, this will then give us the required result.

First we shall show an auxiliary result on the relationship between the cones $\mathcal{K}_n^0$ and $\mathcal{K}_n^r$ for $r \geq 1$.

Lemma 4.1. *Let $n \geq 1$ and $r \geq 0$ be integers. Then*

$$\{A \in \mathbb{S}^n \mid DAD \in \mathcal{K}_n^r \text{ for all } D \in \mathcal{D}\} = \mathcal{K}_n^0.$$

Proof. Since the cone $\mathcal{K}_n^0 = \mathcal{S}_+^n + \mathcal{N}^n$ is invariant under arbitrary scalings, we have that $A \in \mathcal{K}_n^0$ implies $DAD \in \mathcal{K}_n^0 \subset \mathcal{K}_n^r$ for all $D \in \mathcal{D}$. This proves one inclusion, and the whole statement for $r = 0$.

We now prove the other inclusion for $r \geq 1$. Let $A \in \mathbb{S}^n$ be such that $DAD \in \mathcal{K}_n^r$ for all $D \in \mathcal{D}$. Then for all $d_1, \ldots, d_n > 0$, the polynomial

$$\left(\sum_{i,j=1}^{n} A_{ij} d_i d_j x_i^2 x_j^2\right)\left(\sum_{i=1}^{n} x_i^2\right)^r$$

is a sum of squares of polynomials in the variables $x_1, \ldots, x_n$.

Equivalently, $\left(\sum_{i,j=1}^{n} A_{ij} z_i^2 z_j^2\right)\left(\sum_{i=1}^{n} d_i^{-1} z_i^2\right)^r$ is a sum of squares of polynomials in the variables $z_i = \sqrt{d_i} x_i$, $i = 1, \ldots, n$. Let us now fix $d_1 = 1$ and let $d_i \to +\infty$ for $i > 1$. Since the cone of sums of squares polynomials is closed (this result is attributed to Robinson [Rob73]; a more accessible reference where a proof can be found is [Lau09, Section 3.8]), the limit polynomial $\left(\sum_{i,j=1}^{n} A_{ij} z_i^2 z_j^2\right)\left(z_1^2\right)^r$ must also be a sum of squares of polynomials in $z_1, \ldots, z_n$, say $\left(\sum_{i,j=1}^{n} A_{ij} z_i^2 z_j^2\right) z_1^{2r} = \sum_{k=1}^{N} q_k^2(\mathbf{z})$. But then for all k we must have $q_k(\mathbf{z}) = 0$ whenever $z_1 = 0$. It follows that z_1 can be factored out of q_k, i.e., $\left(\sum_{i,j=1}^{n} A_{ij} z_i^2 z_j^2\right) z_1^{2(r-1)}$ is also a sum of squares. After repeatedly carrying out this factoring out process, we arrive at the conclusion that $\left(\sum_{i,j=1}^{n} A_{ij} z_i^2 z_j^2\right)$ is a sum of squares, i.e., $A \in \mathcal{K}_n^0$. This concludes the proof. □

Theorem 4.2. *For any matrix $A \in \mathcal{COP}^n \setminus (\mathcal{S}_+^n + \mathcal{N}^n)$ and any $r \geq 0$, there exists $D \in \mathcal{D}$ such that $DAD \in \mathcal{COP}^n \setminus \mathcal{K}_n^r$.*

Proof. For any $A \in \mathcal{COP}^n$ and $D \in \mathcal{D}$ we have that $DAD \in \mathcal{COP}^n$. Therefore we need only show that for any $A \in \mathcal{COP}^n \setminus (\mathcal{S}_+^n + \mathcal{N}^n)$ there exists $D \in \mathcal{D}$ such that $DAD \notin \mathcal{K}_n^r$.
Assume that such a D does not exist. Then for all $D \in \mathcal{D}$ we have $DAD \in \mathcal{K}_n^r$, and by Lemma 4.1 it follows that $A \in \mathcal{K}_n^0 = \mathcal{S}_+^n + \mathcal{N}^n$, a contradiction. □

Corollary 4.3. *Let $r \geq 0$ be an integer. Then $\mathcal{COP}^n = \mathcal{K}_n^r$ if and only if $n \leq 4$.*

Although we now know from Theorem 4.2 that for every matrix $A \in \mathcal{COP}^n \setminus \mathcal{K}_n^0$ there exists a scaling matrix in $\mathcal{D}$ that scales A out of any cone $\mathcal{K}_n^r$ for $r \geq 1$, $n \geq 5$, finding such a scaling matrix for an arbitrary matrix A is not obvious. For practical purposes however it can be interesting to be able to construct such explicit examples. A result in this direction that generates scalings that scale matrices in $\mathcal{COP}^n \setminus \mathcal{K}_n^0$ out of $\mathcal{K}_n^1$ is presented next. This result comes from correspondence with Peter J.C. Dickinson.

Theorem 4.4. *Let $A \in \mathcal{COP}^n \setminus (\mathcal{S}_+^n + \mathcal{N}^n)$ and let $V^1 \in \mathrm{int}(\mathcal{S}_+^n \cap \mathcal{N}^n)$ be such that $\langle A, V^1 \rangle < 0$. For $i = 2, \dots, n$ define*

$$V^i = \sum_{l=2}^{n} \left(\sqrt{\frac{(V^1)_{1i}}{n-1}}\, \mathbf{e}_1 + \sqrt{\frac{n-1}{(V^1)_{1i}}}\, (V^1)_{il}\mathbf{e}_l \right) \left(\sqrt{\frac{(V^1)_{1i}}{n-1}}\, \mathbf{e}_1 + \sqrt{\frac{n-1}{(V^1)_{1i}}}\, (V^1)_{il}\mathbf{e}_l \right)^{\mathsf{T}}$$
$$+ \sum_{\substack{l=2:\\ l\neq i}}^{n} \left(\mathbf{e}_l + \left(\frac{n-1}{(V^1)_{1l}} (V^1)_{li}^2 + 1 \right) \mathbf{e}_i \right) \left(\mathbf{e}_l + \left(\frac{n-1}{(V^1)_{1l}} (V^1)_{li}^2 + 1 \right) \mathbf{e}_i \right)^{\mathsf{T}}. \tag{4.2}$$

Then for any $D \in \mathcal{D}$ such that $\sum_{i=1}^n \frac{1}{D_{ii}} \langle A, V^i \rangle < 0$ we have $DAD \in \mathcal{COP}^n \setminus \mathcal{K}_n^1$.

Proof. For any $A \in \mathcal{COP}^n$ and $\bar{D} \in \mathcal{D}$ we have that $\bar{D}A\bar{D} \in \mathcal{COP}^n$. Therefore we need only show that for our particular choice of D we have $DAD \notin \mathcal{K}_n^1$. A certificate for $A \notin \mathcal{K}_n^1$ would be a matrix $Y \in \mathcal{K}_n^{1*}$ such that $\langle A, Y \rangle < 0$ and we recall from [Don13] that

$$\mathcal{K}_n^{1*} = \left\{ \sum_{i=1}^{n} Y^i \;\middle|\; Y^i \in \mathcal{S}_+^n \cap \mathcal{N}^n,\ (Y^i)_{jk} = (Y^j)_{ki} \text{ for all } i, j, k = 1, \dots, n \right\}.$$

As $V^1 > 0$ the matrices $V^2, \dots, V^n$ are well defined and we can immediately see that $V^i \in \mathcal{S}_+^n \cap \mathcal{N}^n$ for all i.

Next we show that $(V^i)_{jk} = (V^j)_{ki}$ for all i, j, k in order to show that in fact we have $\sum_{i=1}^n V^i \in \mathcal{K}_n^{1*}$. These equalities can equivalently be written as

$$\begin{aligned} (V^k)_{ij} = (V^i)_{jk} = (V^j)_{ki} \qquad & \text{for all } i, j, k = 1, \dots, n \text{ s.t. } i < j < k, \\ (V^i)_{jj} = (V^j)_{ji} \qquad & \text{for all } i, j = 1, \dots, n \text{ s.t. } i \neq j. \end{aligned}$$

We can now split up the conditions in to the following five cases:

1. For $i, j, k = 2, \dots, n$ with $i < j < k$ we have that $0 = (V^k)_{ij} = (V^i)_{jk} = (V^j)_{ki}$.

2. For $i = 1$ and $j, k = 2, \dots, n$ with $i < j < k$ we have that $(V^k)_{1j} = (V^1)_{jk} = (V^j)_{k1}$.

3. For $i, j = 2, \dots, n$ such that $i \neq j$ we have that

$$(V^i)_{jj} = \frac{n-1}{(V^1)_{1i}} (V^1)_{ij}^2 + 1 = (V^j)_{ji}$$

4. For $j = 1$ and $i = 2, \ldots, n$ we have that
$$(V^i)_{11} = \sum_{l=2}^{n} \frac{(V^1)_{1i}}{n-1} = (V^1)_{1i}.$$

5. For $i = 1$ and $j = 2, \ldots, n$ we have that $(V^j)_{j1} = (V^1)_{jj}$.

We are now ready to prove the proposed scaling out result. Define the matrices
$$Y^i = \frac{1}{D_{ii}} D^{-1} V^i D^{-1} \text{ for all } i = 1, \ldots, n.$$
It follows from the discussion above that for all $i, j, k = 1, \ldots, n$ we have that $Y^i \in \mathcal{S}_+^n \cap \mathcal{N}^n$ and
$$(Y^i)_{jk} = \frac{1}{D_{ii}} \frac{1}{D_{jj}} (V^i)_{jk} \frac{1}{D_{kk}} = \frac{1}{D_{jj}} \frac{1}{D_{kk}} (V^j)_{ki} \frac{1}{D_{ii}} = (Y^j)_{ki}.$$
Therefore we have that $\sum_{i=1}^n Y^i \in \mathcal{K}_n^{1*}$ and
$$\left\langle DAD, \sum_{i=1}^{n} Y^i \right\rangle = \sum_{i=1}^{n} \frac{1}{D_{ii}} \langle A, V^i \rangle < 0,$$
which shows that $DAD \notin \mathcal{K}_n^1$ as claimed. □

We will now illustrate this result with an example that scales the Horn matrix (1.10) out of $\mathcal{K}_5^1$ (recall that we know the Horn matrix to be in $\mathcal{K}_5^1$).

Example 4.5. Consider the Horn matrix from (1.10). Next, let
$$V^1 = \begin{pmatrix} 18 & 12 & 2\frac{1}{2} & 2\frac{1}{2} & 12 \\ 12 & 18 & 12 & 2\frac{1}{2} & 2\frac{1}{2} \\ 2\frac{1}{2} & 12 & 18 & 12 & 2\frac{1}{2} \\ 2\frac{1}{2} & 2\frac{1}{2} & 12 & 18 & 12 \\ 12 & 2\frac{1}{2} & 2\frac{1}{2} & 12 & 18 \end{pmatrix} \in \operatorname{int}(\mathcal{S}_+^5 \cap \mathcal{N}^5).$$
It is easily verified that $\langle H, V^1 \rangle = -5 < 0$. Then via 4.2 we obtain the following inequality:
$$\sum_{i=1}^{5} \frac{1}{D_{ii}} \langle H, V^i \rangle \approx \frac{-5}{D_{11}} + \frac{53405}{D_{22}} + \frac{56428}{D_{33}} + \frac{56428}{D_{44}} - \frac{53405}{D_{55}} < 0. \qquad (4.3)$$
Note that we have an infinite number of scalings that scale the Horn matrix out of $\mathcal{K}_5^1$. One such scaling can be obtained by setting $D_{22} = D_{33} = D_{44} = D_{55} = 1$ and $D_{11} = \frac{1}{50000}$, which clearly satisfies (4.3). We then obtain a matrix DHD which is in $\mathcal{COP}^5$ by Proposition 1.3, but for which it can easily be verified that it is not in $\mathcal{K}_n^1$ using the LMIs (4.1a) - (4.1d).

4.2 Non-decreasing scalings

An immediate question following the result of Theorem 4.2 is whether or not this result also implies the reverse. That is, given $r > s \geq 1$ and a matrix $A \in \mathcal{K}_n^r \setminus \mathcal{K}_n^s$ does there always exist a scaling matrix $D \in \mathcal{D}$ such that $DAD \in \mathcal{K}_n^s$? At the moment of writing this is still an open question, however we can state the result that if such scalings were to exist then the problem of finding them would have to be NP-hard. To prove this result we will initially direct out focus on the cones $\mathcal{C}_n^r$ (1.12) instead. Showing the claimed complexity result for the cones $\mathcal{C}_n^r$ then immediately implies the stated result for the cones $\mathcal{K}_n^r$. We begin by defining the following hypothetical algorithm.

Definition 4.6. Let $r > s \geq 1$. We denote by $\Gamma(\bullet)$ an algorithm that given a matrix $A \in \mathcal{C}_n^r \setminus \mathcal{C}_n^s$ returns a scaling matrix $D \in \mathcal{D}$ such that $DAD \in \mathcal{C}_n^s$.

The proof that we present in this section for the claimed complexity result will go by the means of establishing a polynomial time Turing reduction from the stable set problem to the problem of constructing scaling matrices as produced by $\Gamma(\bullet)$. In particular we will show that if there exists a polynomial time algorithm $\Gamma(\bullet)$ then we can solve the stable set problem in polynomial time.

Note that if such an algorithm $\Gamma(\bullet)$ were to exist, regardless of its complexity, it would imply that the interior of $\mathcal{COP}^n$ reduces to $\mathcal{K}_n^1$ under some scalings $\mathcal{D}^{'} \subseteq \mathcal{D}$, as $D_1 D_2$ is again a diagonal matrix for $D_1, D_2 \in \mathcal{D}$.

Next, recall from [dKP02] that the stability number can be formulated as a copositive program via (1.30). Furthermore recall from [dKP02] the following result regarding relaxations of (1.30) using the hierarchy of cones $\mathcal{C}_n^r$.

Theorem 4.7 (Theorem 4.1, [dKP02])**.** *Let G be a graph with stability number α_G, and let $\zeta^{(r)} := \min\{\lambda \mid (I + A_G)\lambda - E \in \mathcal{C}_n^r\}$, $r \in \mathbb{N}$. Then*

$$\zeta^{(0)} \geq \zeta^{(1)} \geq \ldots \geq \lfloor \zeta^{(r)} \rfloor = \alpha_G$$

for $r \geq \alpha_G^2$.

This result implies that in order to obtain the stability number of a graph it is sufficient to optimize over the polyhedral cone $\mathcal{C}_n^{\alpha_G^2}$. Note that we do need to round when using such relaxations over $\mathcal{C}_n^r$, which means we cannot conclude from Theorem 4.7 that $(I + A_G)\alpha_G - E \in \mathcal{C}_n^{\alpha_G^2}$. However, from [dKP02] we also have the following result.

Lemma 4.8. *Let G be a graph on n vertices with stability number α_G and adjacency matrix A_G. For $\lambda \in \mathbb{R}$ and $\varepsilon = \frac{1}{\lambda + 1/(\lambda-1)} > 0$ let*

$$Q_\lambda = (1+\varepsilon)\lambda(I + A_G) - E. \tag{4.4}$$

Then $Q_{\alpha_G} \in \mathcal{C}_n^{\alpha_G^2}$, while $(1+\varepsilon)\alpha_G < 1 + \alpha_G$.

We now introduce an algorithm that, using $\Gamma(\bullet)$ as a subroutine, decides whether or not the stability number for some graph G is larger or equal to a given $\lambda \in \mathbb{N}$.

Algorithm 1 $F_1(\lambda, A_G)$

Require: $\lambda \in \mathbb{N}$ and A_G

$i \leftarrow 0$
$\varepsilon \leftarrow \frac{1}{\lambda+1/(\lambda-1)}$
$A^{(i)} \leftarrow (1+\varepsilon)(I + A_G)\lambda - E$
while $A^{(i)} \notin \mathcal{C}_n^1$ **and** $i \leq n^2$ **do**
  $D^{(i)} \leftarrow \Gamma(A^{(i)})$
  $A^{(i+1)} \leftarrow D^{(i)} A^{(i)} D^{(i)}$
  $i \leftarrow i + 1$
end while
if $i = n^2$ **then**
  return 0
else
  return 1
end if

We present the following lemma showing that Algorithm 1 does exactly what we claim it to do.

Lemma 4.9. *Consider a graph G on n vertices. Then for any $\lambda \in \mathbb{N}$, Algorithm 1 decides whether or not $\alpha_G \leq \lambda$, that is $F_1(\lambda, A_G) = 1$ if and only if $\alpha_G \leq \lambda$.*

Proof. First, let $\lambda = \alpha_G$. Then from Lemma 4.8 we know that $A^{(0)} \in \mathcal{C}_n^{\alpha_G^2}$. Furthermore, by definition of the algorithm $\Gamma(\bullet)$ we know that there exists a $0 \leq k \leq \alpha_G^2 - 1 < n^2$ such that $A^{(k)} \in \mathcal{C}_n^1$ and hence $F_1(\lambda, A_G) = 1$.
Next, let $\lambda > \alpha_G$, so that $\lambda = \alpha_G + \beta$ for some $\beta > 0$. Then

$$\begin{aligned} Q_\lambda &= (1+\varepsilon)(\alpha_G + \beta)(I + A_G) - E \\ &= (1+\varepsilon)\alpha_G(I + A_G) - E + (1+\varepsilon)\beta(I + A_G) \\ &= Q_{\alpha_G} + (1+\varepsilon)\beta(I + A_G). \end{aligned}$$

Then because $I + A_G \in \mathcal{C}_n^0 \subseteq \mathcal{C}_n^{\alpha_G^2}$, we get $Q_\lambda \in \mathcal{C}_n^{\alpha_G^2}$. The fact that $F_1(\lambda, A_G) = 1$ now follows from the case $\lambda = \alpha_G$.

Finally, let $\lambda < \alpha_G$. Due to the fact that ε is increasing with λ and because $(1+\varepsilon)\alpha_G < 1+\alpha_G$ (see Lemma 4.8), the formulation of the stability number (1.30) implies that $Q_\lambda \notin \mathcal{COP}^n$. Then, because $\mathcal{COP}^n$ is closed under scaling it can never be the case that $DQ_\lambda D \in \mathcal{COP}^n \supseteq \mathcal{C}_n^1$. As a result the **WHILE** loop will conclude only when the condition $i = n^2$ is met and hence the algorithm returns 0, or $F(\lambda, A_G) = 0$. This proves the lemma. □

Next, we demonstrate that we can use Algorithm 1 to determine the stability number of a graph, establishing a Turing reduction as claimed at the start of this section. In particular we define the following algorithm that uses Algorithm 1 as a subroutine:

Algorithm 2 $F_2(A_G)$

Require: A_G

$i \leftarrow 0$
$T \leftarrow 1$
while $T = 1$ **do**
    $T \leftarrow F_1(n - i, A_G)$
    $i \leftarrow i + 1$
end while
return $n - i + 1$

Lemma 4.10. *Consider a graph G on n vertices with adjacency matrix A_G. Then $F_2(A_G) = \alpha_G$.*

Proof. This follows directly from Lemma 4.9 by noting that the first ever instance where $F_1(n - i, A_G)$ returns 0 is when $n - i = \alpha_G - 1$. □

Using the above presented lemmas we can now prove the complexity result stated at the beginning of this section. That is, we show that the Turing reduction established by Lemma 4.10 is a polynomial time Turing reduction from the stable set problem to the problem of finding scalings as produced by the hypothetical algorithm $\Gamma(\bullet)$.

Theorem 4.11. *For $r > s \geq 1$, assume that for any matrix $A \in \mathcal{C}_n^r \setminus \mathcal{C}_n^s$ there exists at least one matrix $D \in \mathcal{D}$ such that $DAD \in \mathcal{C}_n^s$. Then the problem of finding such scaling matrices $D \in \mathcal{D}$ for arbitrary matrices $A \in \mathcal{C}_n^r \setminus \mathcal{C}_n^s$ is NP-hard.*

Proof. Let G be a graph on n vertices and assume that there exists an algorithm $\Gamma(\bullet)$ as in Definition 4.6. Furthermore assume that this is a polynomial time algorithm. We will now show that the existence of such a polynomial time algorithm together with Algorithm 2 establishes a method to compute the stability number in polynomial time, which would prove our result.

First, it should be noted that by definition the encoding lengths of the entries on the diagonal of any scaling matrix returned by $\Gamma(\bullet)$ are polynomially bounded. If this were not the case it would contradict the assumption that $\Gamma(\bullet)$ is a polynomial time algorithm. Next, the entries of the matrix

$$Q_\lambda = (1+\varepsilon)(I + A_G)\lambda - E$$

with $\varepsilon = \frac{1}{\lambda+1/(\lambda-1)}$ for $1 \leq \lambda \leq n$ can only have two possible values, $(1+\varepsilon)\lambda$ and -1. Due to the fact that $\lambda \in \mathbb{N}$ and because λ is bounded from above by n it can easily be seen that the encoding lengths for both of these values must be polynomially bounded in n. Finally, this means the encoding lengths of the entries of $D_1 \dots D_k Q_\lambda D_k \dots D_1$ are bounded for all matrices $D_1, \dots, D_k$ returned by the algorithm $\Gamma(\bullet)$. We can now apply Algorithm 2 to determine the clique number of G while being assured that the encoding lengths of all numbers involved are polynomially bounded in n.

Next, observe that the **WHILE** loop in Algorithm 2 is called at most $n - (\alpha_G - 1) \leq n$ times. During every iteration the algorithm calls the subroutine defined by Algorithm 1. The **WHILE** loop of that algorithm has at most n^2 iterations, during each of which the algorithm $\Gamma(\bullet)$ is called, which by the assumption has a time complexity of order $\mathcal{O}(q)$ for some $q \in \mathbb{R}[n]$. Furthermore we need to compute $DA^{(i)}D$ for each pass through of the **WHILE** loop, which requires at most $2n^2$ multiplications due to the fact that D is diagonal. Finally at the end of every iteration we need to check whether or not $A^{(i+1)} \in \mathcal{C}_n^1$ which can be done in polynomial time with respect to n, i.e. has time complexity $\mathcal{O}(p)$ for some $p \in \mathbb{R}[n]$. In total this means that Algorithm 2 has a worst case time complexity of order $\mathcal{O}\big(n(2n^2 + q + p)\big)$ where q and p are both polynomially bounded in n. Together with Theorem 4.10 and the observation that the encoding lengths of all numbers present in Algorithm 2 at any time are polynomially bounded in n, this concludes the proof. □

Theorem 4.11 shows that there is no general computationally easy method that is guaranteed to scale matrices into lower levels of the hierarchy of the polyhedral cones $\mathcal{C}_n^r$ unless P= NP. Moreover, it almost immediately implies the same result for the hierarchy of cones $\mathcal{Q}_n^r$ and in particular $\mathcal{K}_n^r$, as claimed at the start of this section. We summarize this result in the following Corollary.

Corollary 4.12. *Let $\mathbb{Y}_n^r \in \{\mathcal{C}_n^r, \mathcal{Q}_n^r, \mathcal{K}_n^r\}$ and $r > s \geq 1$. Furthermore let $A \in \mathbb{Y}_n^r \setminus \mathbb{Y}_n^s$. Now assume that there exists at least one matrix $D \in \mathcal{D}$ such that $DAD \in \mathbb{Y}_n^s$. Then the problem of finding such a scaling matrix D is NP-hard.*

Proof. follows directly from Theorem 4.11 and the fact that $\mathcal{C}_n^r \subseteq \mathcal{Q}_n^r \subseteq \mathcal{K}_n^r$ for $n \in \mathbb{N}$ and $r \in \mathbb{Z}_+$. $\square$

Note that this result does not imply that scaling cannot be used to move down in any given hierarchy, but rather that efficient algorithms to construct such scalings for arbitrary matrices $A \in \mathcal{C}_n^r$ do not exist. However we might still be able to construct matrices $D \in \mathcal{D}$ which at least do not scale a matrix out of $\mathcal{C}_n^r$ for a given $r \in \mathbb{N}$. Moreover if we choose such a scaling matrix in a clever way there is the possibility that it scales a given matrix downwards into any of the hierarchies. We will call such scalings *non-decreasing scalings*. A trivial example of a non-decreasing scaling matrix is the identity matrix.

Scaling a matrix is computationally cheap, making them interesting object for preprocessing purposes when optimizing over any of the hierarchies mentioned in this section. A relatively simple example of a type of non-decreasing scaling that, unlike the identity matrix, might be useful to apply concerns matrices with a nonnegative row and column.

Theorem 4.13. *Let $A \in \mathcal{COP}^n$ and assume that for some $i \in \{1, \ldots, n\}$ we have $(A)_{ij} \geq 0$ for every $j = 1, \ldots, n$. Then for $D = I + \lambda \mathbf{e}_i \mathbf{e}_i^\mathsf{T}$ with $\lambda \geq 0$ and using the notation from Definition 1.15 we get that*

$$r^*_{\mathbb{Y}_n^r}(DAD) \leq r^*_{\mathbb{Y}_n^r}(A)$$

for $\mathbb{Y}_n^r \in \{\mathcal{C}_n^r, \mathcal{Q}_n^r, \mathcal{K}_n^r\}$.

Proof. Note that by the properties of A and D we have $DAD \geq A$ entry-wise. This immediately implies that the coefficients of the polynomial

$$\Big(\sum_{i=1}^{n} x_i\Big)^r \mathbf{x}^\mathsf{T} DAD\mathbf{x}$$

are greater than or equal to the coefficients of

$$\Big(\sum_{i=1}^{n} x_i\Big)^r \mathbf{x}^\mathsf{T} A\mathbf{x},$$

and hence $r^*_{\mathcal{C}_n^r}(DAD) \leq r^*_{\mathcal{C}_n^r}(A)$. For the hierarchies $\mathcal{Q}_n^r$ and $\mathcal{K}_n^r$, the result follows from the fact that $\mathcal{C}_n^r \subseteq \mathcal{Q}_n^r \subseteq \mathcal{K}_n^r$. $\square$

The idea behind this type of scaling is that the coefficients of the polynomial $f(\mathbf{x}) = (\sum_{i=1}^{n})^r \mathbf{x}^\mathsf{T} A\mathbf{x}$ are formed by linear combinations of the entries of $A \in \mathbb{S}^n$. Therefore, making several entries larger while leaving all the others unchanged could in theory decrease the lifting rank $r^*_{\mathcal{C}_n^r}(A)$. We provide the following example to illustrate this idea.

Example 4.14. Consider the following matrix

$$A = \begin{pmatrix} 2 & 1 & 1 & 1 & 1 \\ 1 & 2 & 1 & 1 & 1 \\ 1 & 1 & 2 & 1 & 1 \\ 1 & 1 & 1 & 5 & -\frac{5}{2} \\ 1 & 1 & 1 & -\frac{5}{2} & 7 \end{pmatrix} \in \mathcal{C}_5^2 \setminus \mathcal{C}_5^1.$$

Then taking $D = I + 9\mathbf{e}_1\mathbf{e}_1^\mathsf{T} + 9\mathbf{e}_2\mathbf{e}_2^\mathsf{T} + 9\mathbf{e}_3\mathbf{e}_3^\mathsf{T}$ it can easily be checked that $DAD \in \mathcal{C}_5^1$.

The case where a row and column contain both positive and negative values appears to be more complicated. Yet, in that case scalings can still be used to go down in the aforementioned hierarchies. A simple example to show this would again be the matrix A given in Example 4.14 by simply multiplying the scaling defined in that example by $1/10$. That way we are technically scaling rows 4 and 5 of A which contain both positive and negative values. An example for a matrix that does not have any nonnegative rows or columns at all can be obtained by simply reversing the scaling that was found for the Horn matrix in Example 4.5.

A related interesting question is whether non-decreasing scalings could be constructed and applied in a generic way to some possibly well structured matrix variables of a copositive program. For example the entries of the matrix $(I + A_G)\lambda - E$ that appears in the copositive formulation of the stability problem (1.30) can be equal to only two different values, $\lambda - 1$ and -1. Moreover we know exactly where the negative entries occur for a given graph G, as well as their value which is constant. If it would be possible to construct non-decreasing scalings for this particular matrix it would open up the possibility of improving the bounds obtained in [dKP02] via $\mathcal{C}_n^r$ and $\mathcal{K}_n^r$, without actually having to go up in the respective hierarchies.

4.3 Scaling a matrix into $\mathcal{K}_5^1$

In this section, we will show that in the 5×5 case it is possible to scale any copositive matrix into $\mathcal{K}_5^1$. More precisely, we show that for any $X \in \mathcal{COP}^5$ there exists a scaling D such that $DXD \in \mathcal{K}_5^1$. To this end, we make use of the theory developed in Chapter 3, where we investigated matrices of the form (3.5), i.e.

$$S(\boldsymbol{\theta}) := \begin{pmatrix} 1 & -\cos\theta_1 & \cos(\theta_1+\theta_2) & \cos(\theta_4+\theta_5) & -\cos\theta_5 \\ -\cos\theta_1 & 1 & -\cos\theta_2 & \cos(\theta_2+\theta_3) & \cos(\theta_5+\theta_1) \\ \cos(\theta_1+\theta_2) & -\cos\theta_2 & 1 & -\cos\theta_3 & \cos(\theta_3+\theta_4) \\ \cos(\theta_4+\theta_5) & \cos(\theta_2+\theta_3) & -\cos\theta_3 & 1 & -\cos\theta_4 \\ -\cos\theta_5 & \cos(\theta_5+\theta_1) & \cos(\theta_3+\theta_4) & -\cos\theta_4 & 1 \end{pmatrix}$$

where $\boldsymbol{\theta} \in \Theta := \left\{\boldsymbol{\theta} \in \mathbb{R}^5_+ \;\middle|\; \mathbf{e}^\mathsf{T}\boldsymbol{\theta} < \pi\right\}$. We will show below in Theorem 4.15 that $S(\boldsymbol{\theta}) \in \mathcal{K}_5^1$. Combining this result with Corollary 3.35 will then immediately imply the proposed scaling result: Take $A \in \mathcal{COP}^5 \setminus (\mathcal{S}^5_+ + \mathcal{N}^5)$. If A has a zero diagonal entry, say $A_{11} = 0$, then copositivity of A implies that its first row and column must be nonnegative, so we can decompose A as

$$A = \begin{pmatrix} 0 & \mathbf{b}^\mathsf{T} \\ \mathbf{b} & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & C \end{pmatrix} \tag{4.5}$$

where $\mathbf{b} \in \mathbb{R}^4$ is nonnegative and $C \in \mathcal{COP}^4 = \mathcal{S}^4_+ + \mathcal{N}^4$, from which we immediately get $A \in \mathcal{K}_5^0 \subseteq \mathcal{K}_5^1$. Now assume $\operatorname{diag}(A) > 0$. Then we use Corollary 3.35 to get $A = DP^\mathsf{T}S(\boldsymbol{\theta})PD + N$. But this is equivalent to

$$D^{-1}AD^{-1} = P^\mathsf{T}S(\boldsymbol{\theta})P + D^{-1}ND^{-1}$$

which is then a decomposition of $D^{-1}AD^{-1}$ as a sum of two elements in $\mathcal{K}_5^1$ (observe that $\mathcal{K}_5^1$ is closed under permutations), so we conclude $D^{-1}AD^{-1} \in \mathcal{K}_5^1$.

Observe that the diagonal entries of $D^{-1}AD^{-1}$ are all equal 1 because of $(S(\boldsymbol{\theta}))_{ii} = 1$ and $(N)_{ii} = 0$. In case there was a zero entry $A_{ii} = 0$, we can analogously scale the submatrix C in (4.5) to 0/1 diagonal. This shows that for any matrix $A \in \mathcal{COP}^5$, scaling A to 0/1 diagonal will yield a matrix in $\mathcal{K}_5^1$.

So for the above arguments to be valid, it remains to show the following theorem.

Theorem 4.15. *For all $\boldsymbol{\theta} \in \Theta$, we have that $S(\boldsymbol{\theta}) \in \mathcal{K}_5^1$.*

Proof. Recall that $S(\boldsymbol{\theta}) \in \mathcal{K}_5^1$ if and only if the following system of LMIs has a feasible solution $M^1(\boldsymbol{\theta}), \dots, M^5(\boldsymbol{\theta})$:

$$S(\boldsymbol{\theta}) - M^i(\boldsymbol{\theta}) \in \mathcal{S}^5_+ \qquad i = 1, \dots, 5 \tag{4.6a}$$

$$(M^i(\boldsymbol{\theta}))_{ii} = 0 \qquad i = 1, \dots, 5 \tag{4.6b}$$

$$(M^i(\boldsymbol{\theta}))_{jj} + 2(M^j(\boldsymbol{\theta}))_{ij} = 0 \qquad i, j = 1, \dots, 5 \quad \text{s.t. } i \neq j \tag{4.6c}$$

$$(M^i(\boldsymbol{\theta}))_{jk} + (M^j(\boldsymbol{\theta}))_{ik} + (M^k(\boldsymbol{\theta}))_{ij} \geq 0 \quad i, j, k = 1, \dots, 5 \quad \text{s.t. } i < j < k. \tag{4.6d}$$

We claim that the following matrices constitute a feasible solution for this system:

$$M^1(\boldsymbol{\theta}) = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \alpha_1(\boldsymbol{\theta}) & 0 \\ 0 & 0 & 0 & \beta_1(\boldsymbol{\theta}) & \gamma_1(\boldsymbol{\theta}) \\ 0 & \alpha_1(\boldsymbol{\theta}) & \beta_1(\boldsymbol{\theta}) & 0 & 0 \\ 0 & 0 & \gamma_1(\boldsymbol{\theta}) & 0 & 0 \end{pmatrix},$$

$$M^2(\boldsymbol{\theta}) = \begin{pmatrix} 0 & 0 & 0 & \gamma_2(\boldsymbol{\theta}) & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \alpha_2(\boldsymbol{\theta}) \\ \gamma_2(\boldsymbol{\theta}) & 0 & 0 & 0 & \beta_2(\boldsymbol{\theta}) \\ 0 & 0 & \alpha_2(\boldsymbol{\theta}) & \beta_2(\boldsymbol{\theta}) & 0 \end{pmatrix},$$

$$M^3(\boldsymbol{\theta}) = \begin{pmatrix} 0 & 0 & 0 & \alpha_3(\boldsymbol{\theta}) & \beta_3(\boldsymbol{\theta}) \\ 0 & 0 & 0 & 0 & \gamma_3(\boldsymbol{\theta}) \\ 0 & 0 & 0 & 0 & 0 \\ \alpha_3(\boldsymbol{\theta}) & 0 & 0 & 0 & 0 \\ \beta_3(\boldsymbol{\theta}) & \gamma_3(\boldsymbol{\theta}) & 0 & 0 & 0 \end{pmatrix},$$

$$M^4(\boldsymbol{\theta}) = \begin{pmatrix} 0 & \beta_4(\boldsymbol{\theta}) & \gamma_4(\boldsymbol{\theta}) & 0 & 0 \\ \beta_4(\boldsymbol{\theta}) & 0 & 0 & 0 & \alpha_4(\boldsymbol{\theta}) \\ \gamma_4(\boldsymbol{\theta}) & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & \alpha_4(\boldsymbol{\theta}) & 0 & 0 & 0 \end{pmatrix},$$

$$M^5(\boldsymbol{\theta}) = \begin{pmatrix} 0 & 0 & \alpha_5(\boldsymbol{\theta}) & 0 & 0 \\ 0 & 0 & \beta_5(\boldsymbol{\theta}) & \gamma_5(\boldsymbol{\theta}) & 0 \\ \alpha_5(\boldsymbol{\theta}) & \beta_5(\boldsymbol{\theta}) & 0 & 0 & 0 \\ 0 & \gamma_5(\boldsymbol{\theta}) & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

where for all $i = 1, \dots, 5$ (the indices being modulo 5) we define

$$\begin{aligned} \alpha_i(\boldsymbol{\theta}) &= \cos(\theta_{i-2} + \theta_{i-1} + \theta_i) + \cos(\theta_{i+1} + \theta_{i+2}), \\ \beta_i(\boldsymbol{\theta}) &= -\cos(\theta_{i+2}) - \cos(\theta_{i-2} + \theta_{i-1} + \theta_i + \theta_{i+1}), \\ \gamma_i(\boldsymbol{\theta}) &= \cos(\theta_{i-1} + \theta_i + \theta_{i+1}) + \cos(\theta_{i+2} + \theta_{i-2}). \end{aligned}$$

From the fact that $(M^i(\boldsymbol{\theta}))_{jj} = (M^i(\boldsymbol{\theta}))_{ij} = 0$ for all $i, j = 1, \dots, 5$ it is immediately clear that (4.6b) and (4.6c) hold for $M^1(\boldsymbol{\theta}), \dots, M^5(\boldsymbol{\theta})$.

To show (4.6a) we first note that due to the cyclic symmetry in $S(\boldsymbol{\theta})$ and the $M^i(\boldsymbol{\theta})$'s we need only prove this for a single i. Taking $i = 1$ we have the following factorization, where we use the well known formula that $\cos(a+b) = \cos a \cos b - \sin a \sin b$:

$$S(\boldsymbol{\theta}) - M^1(\boldsymbol{\theta})$$

$$= \begin{pmatrix} 1 & -\cos\theta_1 & \cos(\theta_1+\theta_2) & \cos(\theta_4+\theta_5) & -\cos\theta_5 \\ -\cos\theta_1 & 1 & -\cos\theta_2 & -\cos(\theta_4+\theta_5+\theta_1) & \cos(\theta_5+\theta_1) \\ \cos(\theta_1+\theta_2) & -\cos\theta_2 & 1 & \cos(\theta_4+\theta_5+\theta_1+\theta_2) & -\cos(\theta_5+\theta_1+\theta_2) \\ \cos(\theta_4+\theta_5) & -\cos(\theta_4+\theta_5+\theta_1) & \cos(\theta_4+\theta_5+\theta_1+\theta_2) & 1 & -\cos\theta_4 \\ -\cos\theta_5 & \cos(\theta_5+\theta_1) & -\cos(\theta_5+\theta_1+\theta_2) & -\cos\theta_4 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 1 \\ -\cos\theta_1 \\ \cos(\theta_1+\theta_2) \\ \cos(\theta_4+\theta_5) \\ -\cos\theta_5 \end{pmatrix} \begin{pmatrix} 1 \\ -\cos\theta_1 \\ \cos(\theta_1+\theta_2) \\ \cos(\theta_4+\theta_5) \\ -\cos\theta_5 \end{pmatrix}^{\mathsf{T}} + \begin{pmatrix} 0 \\ \sin\theta_1 \\ -\sin(\theta_1+\theta_2) \\ \sin(\theta_4+\theta_5) \\ -\sin\theta_5 \end{pmatrix} \begin{pmatrix} 0 \\ \sin\theta_1 \\ -\sin(\theta_1+\theta_2) \\ \sin(\theta_4+\theta_5) \\ -\sin\theta_5 \end{pmatrix}^{\mathsf{T}}.$$

Hence $S(\boldsymbol{\theta}) - M^1(\boldsymbol{\theta}) \in \mathcal{S}^5_+$ and (4.6a) is shown.

Showing (4.6d) requires a bit more work. We shall do this by first noting that for most indices this condition is equivalent to $0 \geq 0$, and for the remaining indices it is equivalent to (the indices being modulo 5)

$$\alpha_j(\boldsymbol{\theta}) + \beta_{j-2}(\boldsymbol{\theta}) + \gamma_{j+1}(\boldsymbol{\theta}) \geq 0 \qquad j = 1, \dots, 5, \quad \boldsymbol{\theta} \in \operatorname{cl}\Theta.$$

We recall that for $a, b, c \in \mathbb{R}$ we have that

$$\begin{aligned} \cos a + \cos b &= 2\cos\left(\tfrac{1}{2}(a+b)\right)\cos\left(\tfrac{1}{2}(b-a)\right), \\ \cos a - \cos b &= 2\sin\left(\tfrac{1}{2}(a+b)\right)\sin\left(\tfrac{1}{2}(b-a)\right). \end{aligned}$$

From this we also get that

$$\begin{aligned} &\cos(a+b-c) + \cos(a-b+c) - \cos(-a+b+c) \\ &\quad = \cos(a+b+c) + \Big(\cos(a+b-c) + \cos(a-b+c)\Big) \\ &\qquad\qquad - \Big(\cos(-a+b+c) + \cos(a+b+c)\Big) \\ &\quad = \cos(a+b+c) + 2\cos a \cos(b-c) - 2\cos a\cos(b+c) \\ &\quad = \cos(a+b+c) + 4\cos a \sin b \sin c. \end{aligned}$$

Using these trigonometric identities, we get the following.

$$\begin{aligned}
&\alpha_j(\boldsymbol{\theta}) + \beta_{j-2}(\boldsymbol{\theta}) + \gamma_{j+1}(\boldsymbol{\theta}) \\
&\quad = \cos(\theta_{j-2} + \theta_{j-1} + \theta_j) + \cos(\theta_{j+1} + \theta_{j+2}) - \cos(\theta_j) \\
&\qquad\qquad - \cos(\theta_{j+1} + \theta_{j+2} + \theta_{j-2} + \theta_{j-1}) \\
&\qquad\qquad + \cos(\theta_j + \theta_{j+1} + \theta_{j+2}) + \cos(\theta_{j-2} + \theta_{j-1}) \\
&\quad = 2\cos\left(\tfrac{1}{2}\mathbf{e}^\mathsf{T}\boldsymbol{\theta}\right)\Big[\cos\left(\tfrac{1}{2}(\theta_j + (\theta_{j-2} + \theta_{j-1}) - (\theta_{j+1} + \theta_{j+2}))\right) \\
&\qquad\qquad - \cos\left(\tfrac{1}{2}(-\theta_j + (\theta_{j-2} + \theta_{j-1}) + (\theta_{j+1} + \theta_{j+2}))\right) \\
&\qquad\qquad + \cos\left(\tfrac{1}{2}(\theta_j - (\theta_{j-2} + \theta_{j-1}) + (\theta_{j+1} + \theta_{j+2}))\right)\Big] \\
&\quad = 2\cos\left(\tfrac{1}{2}\mathbf{e}^\mathsf{T}\boldsymbol{\theta}\right)\Big[\cos\left(\tfrac{1}{2}\mathbf{e}^\mathsf{T}\boldsymbol{\theta}\right) \\
&\qquad\qquad + 4\cos\left(\tfrac{1}{2}\theta_j\right)\sin\left(\tfrac{1}{2}(\theta_{j-2} + \theta_{j-1})\right)\sin\left(\tfrac{1}{2}(\theta_{j+1} + \theta_{j+2})\right)\Big].
\end{aligned}$$

This can now clearly be seen to be greater than or equal to zero for all $j = 1, \ldots, 5$, $\boldsymbol{\theta} \in \Theta$. $\square$

We can now summarize the discussions of this section in the following corollary and theorems:

Corollary 4.16. *Let $A \in \mathbb{S}^5$ such that $A_{ii} \in \{0,1\}$ for all i. Then $A \in \mathcal{COP}^5$ if and only if $A \in \mathcal{K}_5^1$.*

This gives us the following interesting result.

Theorem 4.17. *Let $A \in \mathbb{S}^5$ and $D \in \mathcal{D}$ be such that $(DAD)_{ii} \in \{0,1\}$ for all i. Then we have that $A \in \mathcal{COP}^5$ if and only if $DAD \in \mathcal{K}_5^1$.*

This means that if we want to test a matrix $A \in \mathbb{S}^5$ with nonnegative diagonal entries for copositivity, then it suffices to scale it so that its diagonal entries become binary, and to test the scaled matrix for inclusion in $\mathcal{K}_5^1$. The original matrix will be copositive if and only if the scaling is in $\mathcal{K}_5^1$.

Finally we have the following result on a characterization of $\mathcal{COP}^5$.

Theorem 4.18. *For matrices of order five, we have the following characterization of $\mathcal{COP}^5$:*

$$\mathcal{COP}^5 = \{DAD \mid A \in \mathcal{K}_5^1,\ D \in \mathcal{D}\}.$$

We will now briefly illustrate this result with the help of one of our earlier examples.

Example 4.19. Consider the matrix as defined by (**??**). We know that this matrix is in $\mathcal{COP}^5$ but not in $\mathcal{K}_5^1$. Obviously scaling this matrix in such a way that every diagonal entry is equal to 1 simply returns the Horn matrix again which we know to be in $\mathcal{K}_n^1$. In fact such a scaling is exactly the inverse of the scaling D we found in Example 4.5.

4.4 Conjectures and open problems

In this chapter we elaborated some very surprising results, and we are left with some open questions which we will discuss in this section. We will formulate a number of conjectures for which we also provide supporting evidence.

In Section 4.3 we saw that for $n \leq 5$ and $r = 1$ we have that

$$\mathcal{COP}^n = \{DAD \mid A \in \mathcal{K}_n^r,\ D \in \mathcal{D}\}. \tag{4.7}$$

It is an open question as to whether a similar result holds for higher n. It could be that (4.7) holds with $r = 1$ for all $n \geq 1$, or alternatively the weaker statement that for all $n \geq 1$ there exists an $r \geq 0$ such that (4.7) could hold.

In fact in Section 4.3 we found that for $n = 5$ and $r = 1$, it was useful to scale the matrix to have binary values on the diagonal. We extend this with the following conjecture.

Conjecture 4.20. *For all $n \geq 1$, there exists a finite $r \geq 0$ such that for any matrix $A \in \mathbb{S}^n$, with $A_{ii} \geq 0$ for all i, and any $D \in \mathcal{D}$ such that $(DAD)_{ii} \in \{0, 1\}$ for all i we have that $A \in \mathcal{COP}^n$ if and only if $DAD \in \mathcal{K}_n^r$.*

Numerical experiments with randomly generated instances carried out by the authors suggested that if we take a matrix in $\mathcal{K}_n^1$ and scale it such that the diagonal is binary then the scaled version of the matrix would also be in $\mathcal{K}_n^1$. An extension and reinterpretation of this result is the following.

Conjecture 4.21. *Let $A \in \mathbb{S}^n$ such that $A_{ii} \in \{0, 1\}$ for all i. Then we have that $A \notin \mathcal{K}_n^r$ implies that $DAD \notin \mathcal{K}_n^r$ for all $D \in \mathcal{D}$.*

This conjecture effectively says that if you have an arbitrary matrix, and you wish to use one of the Parrilo cones as an approximation for testing if this matrix is copositive, then the best thing to do is to first scale the matrix so that its diagonal entries are binary, as, if this is not in the Parrilo cone, then no scaling of it will be in the Parrilo cone.

These conjectures suggest the importance of scaling the diagonal to binary for the Parrilo cones. One piece of support for this idea for $n \geq 6$ is the following theorem.

Theorem 4.22. *Let $A \in \mathcal{COP}^n$ be such that $A_{ij} \in \{-1, +1\}$ for all $i, j = 1, \ldots, n$. Then we have that $A \in \mathcal{K}_n^1$.*

Proof. We consider an arbitrary $A \in \mathcal{COP}^n$ such that $A_{ij} \in \{-1, +1\}$ for all i, j. First note that we must have that $A_{ii} = 1$ for all i.

We now let $M^1, \ldots, M^n \in \mathbb{S}^n$ be given as follows,

$$(M^i)_{jk} = A_{jk} - A_{ij}A_{ik} \qquad \text{for all } i, j, k = 1, \ldots, n.$$

We claim that these provide a feasible solution to the system of LMIs (4.1a)–(4.1d), and thus a certificate for $A \in \mathcal{K}_n^1$.

From construction it is immediately apparent that for all i we have that $A - M^i$ is a rank 1, positive semidefinite matrix, and so (4.1a) holds.

For all $i, j = 1, \ldots, n$ we have that

$$\begin{aligned} (M^i)_{jj} &= A_{jj} - A_{ij}^2 = 1 - (\pm 1)^2 = 0, \\ (M^i)_{ij} &= A_{ij} - A_{ii}A_{ij} = A_{ij} - A_{ij} = 0. \end{aligned}$$

From this we immediately get that (4.1b) and (4.1c) hold.

We are now left to show that (4.1d) holds. Suppose for the sake of contradiction that there exists an $i < j < k$ such that

$$\begin{aligned} 0 &> (M^i)_{jk} + (M^j)_{ik} + (M^k)_{ij} \\ &= A_{jk} + A_{ik} + A_{ij} - A_{ij}A_{ik} - A_{ij}A_{jk} - A_{ik}A_{jk}. \end{aligned}$$

As all the elements of A are in $\{-1, +1\}$, we must have that $-1 = A_{jk} = A_{ik} = A_{ij}$. However if we now let $\mathbf{z} \in \{0, 1\}^n \subset \mathbb{R}_+^n$ such that

$$(\mathbf{z})_l = \begin{cases} 1 & \text{if } l \in \{i, j, k\} \\ 0 & \text{otherwise,} \end{cases}$$

then we get the contradiction that

$$0 \leq \mathbf{z}^\mathsf{T} A \mathbf{z} = A_{ii} + A_{jj} + A_{kk} + 2A_{ij} + 2A_{ik} + 2A_{jk} = -3 < 0$$

and the proof is complete. □

Copositive matrices with all entries equal to ± 1 were previously studied in [HH69]. We have shown that these matrices must be in $\mathcal{K}_n^1$, even though this includes matrices not in $\mathcal{S}_+^n + \mathcal{N}^n$ (for example the Horn matrix) which from Section 4.1 we know can be scaled out of $\mathcal{K}_n^1$.

Equivalent to scaling to binary would be to scale such that all the diagonal entries are either equal to zero or equal to the same positive scalar. We shall now see that further support for this type of scaling comes from the use of the Parrilo cones in approximating the stability number of a graph. It was shown in [dKP02] that the stability number $\alpha(G)$ of a graph with n nodes and adjacency matrix A_G can be computed as

$$\alpha(G) = \min\{\lambda \in \mathbb{R} \mid \lambda(I + A_G) - E \in \mathcal{COP}^n\},$$

where E is the all ones matrix and I is the identity matrix. An approximation of this can then be provided by replacing the copositive cone with one of the Parrilo cones. This provides a very good approximation in practice.

Since A_G has a zero diagonal we have that for any λ the diagonal entries of $\lambda(I + A_G) - E$ are all equal. This means that the matrix is already scaled in the way that these conjectures would suggest is best, which could give an interpretation as to why the approximations are so good. In fact for the 5×5 case an analogous result of Corollary 4.16 applies. Consequently, our results show that for any graph G with five nodes we have that

$$\begin{aligned} \alpha(G) &= \min\{\lambda \in \mathbb{R} \mid \lambda(I + A_G) - E \in \mathcal{COP}^5\} \\ &= \min\{\lambda \in \mathbb{R} \mid \lambda(I + A_G) - E \in \mathcal{K}_5^1\}. \end{aligned} \tag{4.8}$$

This fact has been observed by De Klerk and Pasechnik [dKP02] for the case where G is the 5-cycle. All other graphs with five vertices are perfect, and it is known [PVZ07, Corollary 15] that the $\mathcal{K}_n^0$-approximation (and hence also the $\mathcal{K}_n^1$-approximation) provides an exact answer for perfect graphs. Therefore, (4.8) is implicitly known, but it seems that it has never been stated explicitly.

Chapter 5

Graph isomorphism and copositive programming

Let $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$ be two graphs with vertex sets V_1 and V_2, and edge sets E_1 and E_2 respectively. Furthermore let $|V_1| = |V_2| = n$. Then the graph isomorphism problem ($\mathcal{GIP}$) is the problem of deciding whether the two graphs are isomorphic, i.e. whether they are the same after a possible relabeling of the vertices. More formally, recalling that we denote by $\mathcal{P}_n$ the set of all permutation matrices, we can define the graph isomorphism problem as follows.

Definition 5.1 (Graph Isomorphism Problem)**.** Given two graphs $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$ with adjacency matrices A and B respectively, the graph isomorphism problem is the problem of deciding whether or not there exists a matrix $X \in \mathcal{P}_n$ such that $A = XBX^\mathsf{T}$. If such a matrix $X \in \mathcal{P}_n$ exists, we say that G_1 is *isomorphic* to G_2.

The complexity of the graph isomorphism problem is currently unknown. That is, it is unknown whether or not the problem is solvable in polynomial time or not. Nor is it known if the $\mathcal{GIP}$ is NP-complete; note that obviously the problem is in NP. This makes the graph isomorphism problem one of the last two problems from the well known list of problems with unknown complexity that was published by Gary and Johnson [GJ79], whose complexity still remains unresolved, the other problem being integer factorization. The problem is generally believed not to be in NP-complete however. Some of the arguments for this belief are the inclusion of $\mathcal{GIP}$ in the class SPP [AK06], and the fact that the polynomial-time hierarchy collapses to its second level if $\mathcal{GIP}$ were to be in NP-complete [Sch88]. This makes the $\mathcal{GIP}$ a possible example of a problem that is neither in P nor NP-complete.
For specific cases the graph isomorphism problem can be solved in polynomial time. Examples include planar graphs [HT71] [HW74], graphs with bounded eigenvalue multiplicity [BGM82] and graphs of bounded degree [Luk82]. For the general case several algorithms have been developed. Some of the most well known are the so called Nauty Algorithm developed by McKay [McK81] and the Weisfeiler-Lehman method [WL68], see [Sch09] for a description of this method in English. The latter is a so called refinement method. It inspects k-tuples of vertices in an iterative manner and assigns attributes to them based on their neighbors' attributes. When this method was first described in the

1970s it was considered a possible solution to the graph isomorphism problem. However in 1992 a family of graphs was constructed by Cai, Fürer, and Immerman in [CFI92] that the Weisfeiler-Lehman method could not distinguish between. That is, it could not correctly decide that the graphs were not the same. It should be noted that technically the Cai, Fürer, and Immerman constructions are examples for graph isomorphism of so called colored graphs. Define, for some graph $G = (V, E)$, the map $\sigma_P : V \to V$ as the function induced by a $P \in \mathcal{P}_n$, that is $\sigma_P(i) = j$ if and only if $P_{ij} = 1$. Furthermore define the map $c_G : V \to U$ that assigns a color to every vertex of G, where U is a set of colors. Then the colored graph isomorphism is defined as follows.

Definition 5.2 (*Colored Graph Isomorphism Problem*)**.** Given two graphs $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$, with adjacency matrices A and B respectively, and some coloring of the vertices c_{G_1} and c_{G_2}, the colored graph isomorphism problem is the problem of deciding whether or not there exists a matrix $X \in \mathcal{P}_n$ such that $A = XBX^\mathsf{T}$ while at the same time $c_{G_1}(v) = c_{G_2}(\sigma_X(v))$ holds for all $v \in V$.

It turns out that the graph isomorphism problem and its colored version are actually equivalent in terms of complexity. A polynomial reduction from the colored to the uncolored graph isomorphism can be found in [Sch09]. In essence, this reduction is done by assigning color-specific attributes to every vertex in the form of rooted trees.

Contrary to the Weisfeiler-Lehman method the Nauty algorithm will always work for every possible pair of graphs. The algorithm does, however, have a theoretically exponential running time. Special instances were constructed by Miyazaki [Miy95], based on the constructions by Cai, Fürer, and Immerman, for which it can be shown that the Nauty Algorithm's running time is exponential. Despite the theoretical worst case running time however, the algorithm is reported in [Sch09] to be able to solve moderately large instances in the order of thousands of vertices for random graphs.

Considering the complexity results as well as the state of the art of current $\mathcal{GIP}$ algorithms, the interest for a copositive formulation comes mostly from the difficult cases, such as the ones constructed by Cai, Fürer and Immerman, rather than the more general case. We will start by defining some of the notation to be used during this chapter. Then we will give a copositive formulation of the $\mathcal{GIP}$ in Section 5.2. We will then provide an LP formulation of the $\mathcal{GIP}$ as well in Section 5.3. In Section 5.4 we will discuss approximations for the $\mathcal{GIP}$, by using known hierarchies of the copositive cone. Moreover we will show that the use of such approximation techniques is appropriate by showing that an answer to the $\mathcal{GIP}$ can always be found for some finite level of these hierarchies. In Section 5.5 we will give an alternative formulation of the $\mathcal{GIP}$

which provides us with a potential approach to establishing polynomial time solvability of the graph isomorphism problem. Finally, during the writing of this chapter we found that another completely positive formulation for the $\mathcal{GIP}$ has been suggested in [MA13]. They consider the completely positive version of the Lovász ϑ function and find that, for any two graphs G_1 and G_2, this function evaluates to n if G_1 is isomorphic to G_2 and less than $n - \frac{1}{4n^4}$ when they are not isomorphic.

5.1 Notation

We recall that the (non-symmetric) matrix E_n^{ij} is defined as an $n \times n$ matrix made up of all zeros, except for the (i, j)-th element, which is equal to one. We denote the Kronecker delta as δ_{ij}. For any matrix A we denote its i^{th} row and column by $\mathbf{A}_{i\bullet}$ and $\mathbf{A}_{\bullet i}$ respectively. Finally, for $X \in \mathcal{P}_n$ we define $\mathbf{x} := \mathrm{Vec}(X)$. During the remainder of this chapter we will slightly abuse this notation by using $\mathbf{x}$ and X interchangeably in an effort to make the text more readable. That is, for example, when we write $\langle D, \mathbf{x}\mathbf{x}^\mathsf{T} \rangle \geq 0$ for all $X \in \mathcal{P}_n$, we mean $\langle D, \mathbf{x}\mathbf{x}^\mathsf{T} \rangle \geq 0$ for every $\mathbf{x} = \mathrm{Vec}(X)$ with $X \in \mathcal{P}_n$. Following the notation used in [PR09], for a matrix $B \in \mathbb{S}^{n^2}$ we use the block notation

$$B = \begin{pmatrix} B^{11} & \dots & B^{1n} \\ \vdots & \ddots & \vdots \\ B^{n1} & \dots & B^{nn} \end{pmatrix}$$

where $B^{ij} \in \mathbb{S}^n$ for every $i, j = 1, \dots, n$. This notation is not to be confused with E_n^{ij}, which will be as it was defined above. In particular, note that we will never use E as anything other than the all ones matrix.

5.2 Graph Isomorphism as a Copositive Program

We will now provide a copositive formulation of the $\mathcal{GIP}$. In order to do so we shall begin by first rewriting the $\mathcal{GIP}$ as a quadratic optimization problem. Once we have done that we can use the technique by Povh and Rendl [PR09] to obtain a copositive program.

We can turn the $\mathcal{GIP}$ into a quadratic optimization problem by rewriting the equality $A = XBX^\mathsf{T}$, from Definition 5.1, into a number of linear constraints as follows,

$$
\begin{aligned}
A = XBX^{\mathsf{T}} &\Leftrightarrow AX = XB \\
&\Leftrightarrow \mathbf{A}_{i\bullet}\mathbf{X}_{\bullet j} - \mathbf{X}_{i\bullet}\mathbf{B}_{\bullet j} = 0 \quad \forall i,j = 1,\ldots,n \\
&\Leftrightarrow \langle (E_n^{ji}A)^{\mathsf{T}}, X\rangle - \langle (BE_n^{ji})^{\mathsf{T}}, X\rangle = 0 \quad \forall i,j = 1,\ldots,n. \\
&\Leftrightarrow \langle (E_n^{ji}A)^{\mathsf{T}} - (BE_n^{ji})^{\mathsf{T}}, X\rangle = 0 \quad \forall i,j = 1,\ldots,n.
\end{aligned}
$$

Note that the matrix E_n^{ij} acts as an operator on a matrix $C \in \mathbb{R}^{n\times n}$ in such a way that $E_n^{ij}C$ is an all zero matrix apart from the i^{th} row, which is identical to the j^{th} row of the matrix C. We can now present the following result

Lemma 5.3 ($\mathcal{GIP}$ as a Quadratic Program)**.** *Given two graphs $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$ with adjacency matrices A and B respectively, define for $i, j = 1, \ldots, n$*

$$
\mathbf{d}_{ij} = \mathrm{Vec}((E_n^{ji}A)^{\mathsf{T}} - (BE_n^{ji})^{\mathsf{T}}) = \mathrm{Vec}(AE_n^{ij} - E_n^{ij}B), \tag{5.1}
$$

and let $D = \sum_{i=1}^{n}\sum_{j=1}^{n}\mathbf{d}_{ij}\mathbf{d}_{ij}^{\mathsf{T}}$.

Then the graph isomorphism problem as described in Definition 5.1 is equivalent to deciding whether or not the optimal value of the quadratic program

$$
\begin{aligned}
\mathcal{GIP}_{QP} = \min \quad & \langle D, \mathbf{x}\mathbf{x}^{\mathsf{T}}\rangle \\
s.t. \quad & X \in \mathcal{P}_n
\end{aligned} \tag{5.2}
$$

is equal to 0.

Proof. From the discussion above, it follows that we can write our equation as follows:

$$
\begin{aligned}
A = XBX^{\mathsf{T}} &\Leftrightarrow \mathbf{d}_{ij}^{\mathsf{T}}\mathbf{x} = 0 && \text{for all } i,j = 1,\ldots,n \\
&\Leftrightarrow (\mathbf{d}_{ij}^{\mathsf{T}}\mathbf{x})^2 = 0 && \text{for all } i,j = 1,\ldots,n \\
&\Leftrightarrow \sum_{i=1}^{n}\sum_{j=1}^{n}(\mathbf{d}_{ij}^{\mathsf{T}}\mathbf{x})^2 = 0 \\
&\Leftrightarrow \sum_{i=1}^{n}\sum_{j=1}^{n}\langle \mathbf{d}_{ij}\mathbf{d}_{ij}^{\mathsf{T}}, \mathbf{x}\mathbf{x}^{\mathsf{T}}\rangle = 0 \\
&\Leftrightarrow \langle D, \mathbf{x}\mathbf{x}^{\mathsf{T}}\rangle = 0.
\end{aligned}
$$

From Proposition 1.2 it follows that $D \in \mathcal{S}_+^{n^2}$ and moreover that $\mathbf{d}_{ij}\mathbf{d}_{ij}^\mathsf{T} \in \mathcal{S}_+^n$ as well. This in turn implies that $\langle D, \mathbf{x}\mathbf{x}^\mathsf{T}\rangle \geq 0$ and $\langle \mathbf{d}_{ij}\mathbf{d}_{ij}^\mathsf{T}, \mathbf{x}\mathbf{x}^\mathsf{T}\rangle \geq 0$ for all $X \in \mathcal{P}_n$ and all $i, j = 1, \ldots, n$, and hence $\langle D, \mathbf{x}\mathbf{x}^\mathsf{T}\rangle = 0$ if and only if $\langle \mathbf{d}_{ij}\mathbf{d}_{ij}^\mathsf{T}, \mathbf{x}\mathbf{x}^\mathsf{T}\rangle = 0$ for every $i, j = 1, \ldots, n$. This proves our claimed result. □

Using the technique from [PR09], as suggested before, we can now write the $\mathcal{GIP}$ as a copositive program. In order to do this we first note that we can write the set of permutation matrices as

$$\mathcal{P}_n = \{A \in \mathbb{R}_+^{n\times n} \mid A^\mathsf{T} A = I_n\}.$$

Then redundant equalities can be added to obtain

$$\mathcal{P}_n = \{A \in \mathbb{R}_+^{n\times n} \mid A^\mathsf{T} A = I_n,\ AA^\mathsf{T} = I_n,\ (\mathbf{e}^\mathsf{T} A\mathbf{e})^2 = n^2\}.$$

Finally, noting that for any matrices $X, B, C \in \mathbb{R}^{n\times n}$ we have $\langle X, BXC\rangle = \langle C^\mathsf{T} \otimes B, \mathbf{x}\mathbf{x}^\mathsf{T}\rangle$, $\mathbf{x} = \mathrm{Vec}(X)$, we obtain the following result using Lagrangian duality, analogue to [PR09]:

$$\begin{aligned}
\mathcal{GIP}_{QP} &= \min\Big\{\langle D, \mathbf{x}\mathbf{x}^\mathsf{T}\rangle \mid X \in \mathcal{P}_n\Big\} \\
&= \min_{X\geq 0}\Big\{\langle D, \mathbf{x}\mathbf{x}^\mathsf{T}\rangle + \min_{S,T\in\mathbb{S}^n, v\in\mathbb{R}}\{\langle S, I - XX^\mathsf{T}\rangle + \langle T, I - X^\mathsf{T} X\rangle + v(n^2 - \langle X, EXE\rangle)\}\Big\} \\
&\geq \max_{S,T\in\mathbb{S}^n, v\in\mathbb{R}}\Big\{\mathrm{Tr}(S) + \mathrm{Tr}(T) + n^2 v + \min_{X\geq 0}\{\mathbf{x}^\mathsf{T}(D - I\otimes S - T\otimes I - vE_{n^2})\mathbf{x}\}\Big\} \\
&= \max_{S,T\in\mathbb{S}^n, v\in\mathbb{R}}\Big\{\mathrm{Tr}(S) + \mathrm{Tr}(T) + n^2 v \mid D - I\otimes S - T\otimes I - vE_{n^2} \in \mathcal{COP}^{n^2}\Big\} \qquad (5.3) \\
&= \min_{Y}\Big\{\langle D, Y\rangle \mid \sum_i Y^{ii} = I,\ \langle I, Y^{ij}\rangle = \delta_{ij} \quad \forall i, j,\ \langle E, Y\rangle = n^2,\ Y \in \mathcal{CP}^{n^2}\Big\} \qquad (5.4)
\end{aligned}$$

The copositive formulation (5.3) will be denoted as $\mathcal{GIP}_{COP}$. That we have strong duality between (5.3) and (5.4) can be seen by taking $T = 0$, $v = 0$, and $S = -I_n$ in the copositive program (5.3), so that the matrix $D + I_{n^2}$ is in the interior of $\mathcal{COP}^{n^2}$. This result follows from the fact that D is positive semidefinite, which as noted before follows directly from its definition and Proposition 1.2. From [PR09] we furthermore get the following result with respect to the feasible region of the completely positive program (5.4).

Theorem 5.4 (Theorem 3, [PR09]). *Let*

$$\mathcal{F} = \Big\{ Y \mid \sum_i Y^{ii} = I,\, \langle I, Y^{ij} \rangle = \delta_{ij}\, \forall i,j,\, \langle E, Y \rangle = n^2,\, Y \in \mathcal{CP}^{n^2} \Big\}, \quad (5.5)$$

then

$$\mathcal{F} = \operatorname{conv}\Big\{ \mathbf{y}\mathbf{y}^\mathsf{T} \mid \mathbf{y} = \operatorname{Vec}(X),\, X \in \mathcal{P}_n \Big\}. \quad (5.6)$$

Hence we now immediately obtain the following corollary.

Corollary 5.5. *The optimal value of* $\mathcal{GIP}_{QP}$ *is equal to the optimal value of* $\mathcal{GIP}_{COP}$.

In other words, we can solve the $\mathcal{GIP}$ by solving a copositive program. We summarize this result in the following theorem.

Theorem 5.6 ($\mathcal{GIP}$ as a Copositive and Completely Positive Program). *Let there be given two graphs* $G_1 = (V_1, E_1)$ *and* $G_2 = (V_2, E_2)$ *with adjacency matrices* A *and* B *respectively. Let* D *be defined as in Lemma 5.3. Then the graph isomorphism as described in Definition 5.1 is equivalent to deciding whether or not the optimal value of the copositive program*

$$\max_{S,T \in \mathbb{S}^n, v \in \mathbb{R}} \Big\{ \operatorname{Tr}(S) + \operatorname{Tr}(T) + n^2 v \mid D - I \otimes S - T \otimes I - vE_{n^2} \in \mathcal{COP}^{n^2} \Big\} \quad (5.7)$$

is equal to 0*, which is the case if and only if* G_1 *and* G_2 *are isomorphic. Alternatively, deciding isomorphism can be done by verifying whether the optimal value of the completely positive program*

$$\min \Big\{ \langle D, Y \rangle \mid \sum_i Y^{ii} = I,\, \langle I, Y^{ij} \rangle = \delta_{ij}\, \forall i,j,\, \langle E, Y \rangle = n^2,\, Y \in \mathcal{CP}^{n^2} \Big\} \quad (5.8)$$

is equal to 0*, which again is the case if and only if* G_1 *is isomorphic to* G_2.

Finally, from Corollary 5.5 we obtain the following lemma that we will need later on in this chapter.

Lemma 5.7. *The optimal values of the copositive formulation* (5.7) *and the completely positive formulation* (5.8) *of the* $\mathcal{GIP}$ *are integer valued.*

Proof. This follows directly from Corollary 5.5 and noting that the matrix D is integer valued. □

5.2.1 Properties of the matrix D

We will now explore some of the properties of the matrix D as defined in Lemma 5.3. We will begin by giving an equivalent definition of D. This alternative definition of D will provide us with a better insight into the structure of this particular matrix. We will then use this insight later on in this chapter to rewrite our current copostive reformulation. In order to do so however, we will first have to recall some technical results concerning the Kronecker product, which we will need in order to be able to obtain a new formulation of our matrix D.

Lemma 5.8 (Proposition 7.1.6, [Ber09]). *Let $A \in \mathbb{R}^{n\times m}$, $B \in \mathbb{R}^{l\times k}$, $C \in \mathbb{R}^{m\times q}$, and $D \in \mathbb{R}^{k\times p}$. Then*

$$(A \otimes B)(C \otimes D) = AC \otimes BD. \tag{5.9}$$

Lemma 5.9 (Proposition 7.1.9, [Ber09]). *Let $A \in \mathbb{R}^{n\times m}$, $B \in \mathbb{R}^{m\times l}$, and $A \in \mathbb{R}^{l\times k}$. Then*

$$\operatorname{Vec}(ABC) = (C^{\mathsf{T}} \otimes A) \operatorname{Vec} B. \tag{5.10}$$

Using the above properties, we can now obtain a more convenient form and definition of our matrix D giving us the following proposition.

Proposition 5.10. *Let $A, B \in \mathbb{S}^n$, and consider a matrix D, defined by $D := \sum_{i=1}^n \sum_{j=1}^n \mathbf{d}_{ij}\mathbf{d}_{ij}^{\mathsf{T}}$, with $\mathbf{d}_{ij} := \operatorname{Vec}(AE_n^{ij} - E_n^{ij}B)$ for every $i, j = 1, \dots, n$. Then*

$$D = (I \otimes AA) + (BB \otimes I) - 2(B \otimes A). \tag{5.11}$$

Proof. Let I be the identity matrix of size $n \times n$ unless otherwise specified. Then we get

$$\begin{aligned}
D :&= \sum_{i=1}^n \sum_{j=1}^n \mathbf{d}_{ij}\mathbf{d}_{ij}^{\mathsf{T}} = \sum_{i=1}^n \sum_{j=1}^n \operatorname{Vec}(AE^{ij} - E^{ij}B) \operatorname{Vec}(AE^{ij} - E^{ij}B)^{\mathsf{T}} \\
&= \sum_{i=1}^n \sum_{j=1}^n \operatorname{Vec}(AE^{ij}) \operatorname{Vec}(AE^{ij})^{\mathsf{T}} + \operatorname{Vec}(E^{ij}B) \operatorname{Vec}(E^{ij}B)^{\mathsf{T}} \\
&\qquad\qquad - \operatorname{Vec}(AE^{ij}) \operatorname{Vec}(E^{ij}B)^{\mathsf{T}} - \operatorname{Vec}(E^{ij}B) \operatorname{Vec}(AE^{ij})^{\mathsf{T}}.
\end{aligned}$$

Treating parts of this summation separately, and using Lemmas 5.8 and 5.9, we obtain the following equations:

$$\sum_{i=1}^{n}\sum_{j=1}^{n}\mathrm{Vec}(AE^{ij})\,\mathrm{Vec}(AE^{ij})^{\mathsf{T}} = \sum_{i=1}^{n}\sum_{j=1}^{n}(I\otimes A)\,\mathrm{Vec}(E^{ij})\Big[(I\otimes A)\,\mathrm{Vec}(E^{ij})\Big]^{\mathsf{T}}$$
$$= (I\otimes A)\Big[\sum_{i=1}^{n}\sum_{j=1}^{n}\mathrm{Vec}(E^{ij})\,\mathrm{Vec}(E^{ij})^{\mathsf{T}}\Big](I\otimes A) = (I\otimes A)I_{n^2}(I\otimes A) = (I\otimes AA),$$

and

$$\sum_{i=1}^{n}\sum_{j=1}^{n}\mathrm{Vec}(E^{ij}B)\,\mathrm{Vec}(E^{ij}B)^{\mathsf{T}} = \sum_{i=1}^{n}\sum_{j=1}^{n}(B\otimes I)\,\mathrm{Vec}(E^{ij})\Big[(B\otimes I)\,\mathrm{Vec}(E^{ij})\Big]^{\mathsf{T}}$$
$$= (B\otimes I)\Big[\sum_{i=1}^{n}\sum_{j=1}^{n}\mathrm{Vec}(E^{ij})\,\mathrm{Vec}(E^{ij})^{\mathsf{T}}\Big](B\otimes I) = (B\otimes I)I_{n^2}(B\otimes I) = (BB\otimes I),$$

and

$$\begin{aligned}(\sum_{i=1}^{n}\sum_{j=1}^{n}\mathrm{Vec}(E^{ij}B)\,\mathrm{Vec}(AE^{ij})^{\mathsf{T}})^{\mathsf{T}} &= \sum_{i=1}^{n}\sum_{j=1}^{n}\mathrm{Vec}(AE^{ij})\,\mathrm{Vec}(E^{ij}B)^{\mathsf{T}}\\ &= \sum_{i=1}^{n}\sum_{j=1}^{n}(I\otimes A)\,\mathrm{Vec}(E^{ij})\Big[(B\otimes I)\,\mathrm{Vec}(E^{ij})\Big]^{\mathsf{T}}\\ &= (I\otimes A)\Big[\sum_{i=1}^{n}\sum_{j=1}^{n}\mathrm{Vec}(E^{ij})\,\mathrm{Vec}(E^{ij})^{\mathsf{T}}\Big](B\otimes I)\\ &= (I\otimes A)I_{n^2}(B\otimes I) = (B\otimes A).\end{aligned}$$

This proves our statement. □

Furthermore recall that due to the original definition of the matrix D we know that it is a sum of semidefinite matrices and hence is itself semidefinite. The fact that $D\in\mathcal{S}_+^{n^2}$ has a number of consequences. One of them is that it places a restriction on the rank of the matrix. Before we can prove this result, recall that from Lemma 3.3 we know that for $A\in\mathcal{S}_+^n$ we have $\mathbf{u}^{\mathsf{T}}A\mathbf{u}=0$ if and only if $A\mathbf{u}=\mathbf{0}$. With the help of this result we can now formulate the following theorem with respect to the rank of D.

Theorem 5.11. *Consider two graphs G_1 and G_2 on n vertices and let A and B be their respective adjacency matrices. Furthermore let*

$$D = (I\otimes AA) + (BB\otimes I) - 2(B\otimes A).$$

If $\operatorname{Rank} D = n^2$, *then G_1 is not isomorphic to G_2.*

Proof. Let $\operatorname{Rank} D = n^2$, then the dimension of the nullspace of D is equal to 0, that is $\operatorname{Null}(D) = \{\mathbf{0}\}$. Now suppose G_1 and G_2 are isomorphic. From Theorem 5.6 it follows that there exists a completely positive matrix $Y \neq 0$ (note that $0 \notin \mathcal{F}$) and nonzero vectors $\mathbf{y}_i \in \mathbb{R}_+^{n^2}$, $i = 1, \dots, k$ for some $k > 0$, such that

$$0 = \langle D, Y \rangle = \langle D, \sum_{i=1}^{k} \mathbf{y}_i \mathbf{y}_i^\mathsf{T} \rangle = \sum_{i=1}^{k} \langle D, \mathbf{y}_i \mathbf{y}_i^\mathsf{T} \rangle.$$

Then because $D \in \mathcal{S}_+^{n^2}$ we know that

$$\langle D, \mathbf{y}_i \mathbf{y}_i^\mathsf{T} \rangle \geq 0 \text{ for all } i.$$

So we must have $\langle D, \mathbf{y}_i \mathbf{y}_i^\mathsf{T} \rangle = 0$ for every i. Then, again because $D \in \mathcal{S}_+^{n^2}$, it follows from Lemma 3.3 that $\mathbf{y}_i^\mathsf{T} D \mathbf{y}_i = 0$ if and only if $D\mathbf{y}_i = 0$. That is, $\mathbf{y}_i$ is in the nullspace of D, which is a contradiction. □

In fact, because we know that a semidefinite matrix is positive definite if and only if it has full rank, we have the following equivalent corollary.

Corollary 5.12. *Consider two graphs G_1 and G_2 on n vertices, and let A and B be their respective adjacency matrices. Furthermore let*

$$D = I \otimes AA + BB \otimes I - 2B \otimes A.$$

If D is positive definite, then G_1 is not isomorphic to G_2.

The reverse does not hold, the following pair of graphs is a counter example.

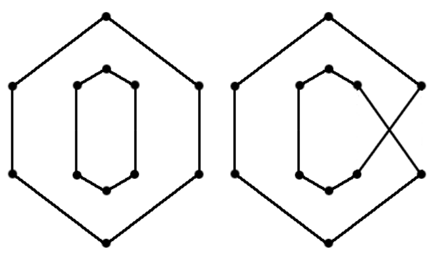

Figure 5.1: *Example of a pair of non-isomorphic graphs created following the construction of Cai, Fürer, and Immerman.*

The resulting matrix D has a rank of 124 rather than $12^2 = 144$. This pair of graphs is an example of the Cai, Fürer, and Immerman constructions that were mentioned earlier in this chapter. Moreover it is a special case as it does not need a coloring for the two graphs to be non-isomorphic. This means we do not have to add rooted graphs to all the vertices in order to do computations on these graphs in an effort to decide isomorphism. As it turns out, we can already find a certificate for non-isomorphism of these two graphs by replacing the copositive cone $\mathcal{COP}^n$ by the semidefinite cone $\mathcal{S}_+^n$ in (5.7). This fact is not necessarily surprising as these graphs are planar graphs, for which we know the $\mathcal{GIP}$ is decidable in polynomial time from [HT71] [HW74], as mentioned at the start of this chapter.

5.3 The Graph Isomorphism Problem as an LP

From Theorem 5.4 we know that we can write our completely positive program (5.8) as follows,

$$\min \left\{ \langle D, Y \rangle \mid \; Y \in \operatorname{conv}\{\mathbf{x}\mathbf{x}^\mathsf{T},\, \mathbf{x} = \operatorname{Vec}(X),\, X \in \mathcal{P}_n\} \right\},$$

or equivalently,

$$\min \left\{ \langle D, Y \rangle \mid \; Y = \sum_{X \in \mathcal{P}_n} \lambda_X \operatorname{Vec}(X) \operatorname{Vec}(X)^\mathsf{T}, \text{ for } \lambda_X \geq 0 \text{ and } \sum_{X \in \mathcal{P}_n} \lambda_X = 1 \right\}.$$

This provides us with a linear programming formulation of the graph isomorphism problem. Note that this does not solve the complexity of the graph isomorphism problem, as we have an exponential number of variables λ_X. By dualizing we obtain an LP with just one variable, but with an exponential number of constraints.

Let $\mathcal{L}(\lambda_X, \mu)$ be the Lagrangian function belonging to the LP above, then using Lagrangian dualization we get

$$\begin{aligned}
\min_{\lambda_X \geq 0} \max_{\mu \in \mathbb{R}} \mathcal{L}(\lambda_X, \mu) &= \min_{\lambda_X \geq 0} \max_{\mu \in \mathbb{R}} \left\{ \left\langle D, \sum_{X \in \mathcal{P}_n} \lambda_X \operatorname{Vec}(X) \operatorname{Vec}(X)^\mathsf{T} \right\rangle + \mu (1 - \sum_{X \in \mathcal{P}_n} \lambda_X) \right\} \\
&\geq \max_{\mu \in \mathbb{R}} \min_{\lambda_X \geq 0} \left\{ \mu (1 - \sum_{X \in \mathcal{P}_n} \lambda_X) + \left\langle D, \sum_{X \in \mathcal{P}_n} \lambda_X \operatorname{Vec}(X) \operatorname{Vec}(X)^\mathsf{T} \right\rangle \right\} \\
&= \max_{\mu \in \mathbb{R}} \left\{ \mu + \min_{\lambda_X \geq 0} \left\{ \sum_{X \in \mathcal{P}_n} \lambda_X \left(\left\langle D, \operatorname{Vec}(X) \operatorname{Vec}(X)^\mathsf{T} \right\rangle - \mu \right) \right\} \right\} \\
&= \max \left\{ \mu \in \mathbb{R} \mid \; \langle D, \operatorname{Vec}(X) \operatorname{Vec}(X)^\mathsf{T} \rangle - \mu \geq 0 \text{ for all } X \in \mathcal{P}_n \right\}. \quad (5.12)
\end{aligned}$$

In the case where we have two isomorphic graphs, D is such that there exists an $X \in \mathcal{P}_n$ for which $\langle D, \mathrm{Vec}(X)\,\mathrm{Vec}(X)^\mathsf{T}\rangle = 0$ by Lemma 5.3, and hence the optimal value for the dual is 0. In the case that our two graphs are not isomorphic we have $\langle D, \mathrm{Vec}(X)\,\mathrm{Vec}(X)^\mathsf{T}\rangle > 0$ for every $X \in \mathcal{P}_n$. However, due to the fact that $-2 \leq D_{ij} \leq 2n-2$ by Proposition 5.10, the inner product $\langle D, \mathrm{Vec}(X)\,\mathrm{Vec}(X)^\mathsf{T}\rangle$ must be finite. Hence μ, and therefore the optimal value of (5.12) must be finite as well. From duality theory of linear programs it now follows that strong duality holds and so the inequality in the above should in fact be an equality. This then provides us with yet another linear programming formulation for the graph isomorphism problem as claimed. We summarize this result in the following theorem.

Theorem 5.13 ($\mathcal{GIP}$ as a Linear Program)**.** *Given two graphs $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$ with adjacency matrices A and B respectively. Let D be defined as in Lemma 5.3. Then the graph isomorphism problem as described in Definition 5.1 is equivalent to deciding whether or not the optimal value of the linear program*

$$\min\Big\{\langle D, Y\rangle \mid \; Y = \sum_{X\in\mathcal{P}_n} \lambda_X\,\mathrm{Vec}(X)\,\mathrm{Vec}(X)^\mathsf{T},\ \textit{for } \lambda_X \geq 0 \textit{ and } \sum_{X\in\mathcal{P}_n} \lambda_X = 1\Big\}$$

$$= \max\Big\{\mu \in \mathbb{R} \mid \; \langle D, \mathrm{Vec}(X)\,\mathrm{Vec}(X)^\mathsf{T}\rangle - \mu \geq 0,\ \textit{for all } X \in \mathcal{P}_n\Big\}.$$

is equal to 0*, which is the case if and only if G_1 and G_2 are isomorphic.*

5.4 Solving the $\mathcal{GIP}$ via approximation hierarchies

We will start this section by recalling the definition of the hierarchy of approximations (1.12) as introduced in [dKP02]. That is for $r \in \mathbb{Z}_+$ and $n \in \mathbb{N}$

$$\mathcal{C}_n^r := \Big\{A \in \mathbb{S}^n \mid \Big(\sum_{i=1}^n x_i\Big)^r \mathbf{x}^\mathsf{T} A\mathbf{x} \text{ has nonnegative coefficients }\Big\}.$$

We then state the following definition.

Definition 5.14. For any two graphs G_1 and G_2 with adjacency matrices A and B, let

$$D = (I \otimes AA) + (BB \otimes I) - 2(B \otimes A).$$

Then for any $r \in \mathbb{Z}_+$ we define the following lower bound for (5.7):

$$\eta^r := \max_{S,T \in \mathbb{S}^n, v \in \mathbb{R}} \Big\{ \mathrm{Tr}(S) + \mathrm{Tr}(T) + n^2 v \mid D - I \otimes S - T \otimes I - vE_{n^2} \in \mathcal{C}^r_{n^2} \Big\}. \quad (5.13)$$

It should be noted that the $\mathcal{GIP}$ can only have 'yes' and 'no' as answers. Because of this we are not really interested in approximations per sé as they do not necessarily give us any information. Hence we need to know whether the optimum of the copositive formulation (5.7) is obtained by some finite level of the approximation hierarchy. In order to establish that this is indeed the case we will try to obtain an upper bound on r such that η^r will give us the optimum of (5.7). To this end we present the following known result by de Klerk and Pasechnik [dKP02] which is an addaptation for copositivity of a result by Powers and Reznik [PR01] who constructed a tight upper bound on the value of N in Theorem 1.13.

Corollary 5.15 ([dKP02], Corollary 3.5)**.** *Let $M \in \mathrm{int}(\mathcal{COP}^n)$, then*

$$P^N(\mathbf{z}) = \Big(\sum_{i,j=1}^n M_{ij} z_i z_j \Big) \Big(\sum_{i=1}^n z_i \Big)^N$$

has only nonnegative coefficients if $N > L/\kappa - 2$ where

$$L = \max_{i,j} |M_{ij}| \quad (5.14)$$

and

$$\kappa = \min_{\mathbf{z} \in \Delta} \mathbf{z}^\mathsf{T} M \mathbf{z} \quad (5.15)$$

where Δ denotes the unit simplex.

In particular this corollary is what provides us with a way of obtaining a bound on the number of liftings, r, needed to obtain an optimal solution for GIP_{CP} by solving η^r instead. More importantly, as long as we can show that an optimal solution inside the interior of $\mathcal{COP}^n$ exists, Corollary 5.15 implies that there exists an r for which η^r can decide the graph isomorphism problem. For this purpose we have the following proposition.

Proposition 5.16. *Given two graphs $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$, let D be the matrix as defined in Lemma 5.3. Then there exists an $r \in \mathbb{Z}_+$ such that $\lceil \eta^{(r)} \rceil$ is equal to the optimal value of $\mathcal{GIP}_{COP}$, and hence G_1 is isomorphic to G_2 if and only if $\eta^{(r)} = 0$.*

Proof. Let S^*, T^* and v^* be optimal solutions to the copositive program (5.7), and let $0 < \varepsilon < \frac{1}{n^2}$. Define the matrix

$$Q_\varepsilon^* = D - I \otimes S^* - T^* \otimes I - (v^* - \varepsilon)E_{n^2}.$$

Due to Theorem 1.13, in order to prove the existence of an $r \in \mathbb{Z}_+$ for which the $\mathcal{GIP}$ can be decided by η^r, it suffices to show that $Q_\varepsilon^* \in \text{int}(\mathcal{COP}^{n^2})$ and that Q_ε^* provides an optimal solution.

Obviously $Q_\varepsilon^* \in \mathcal{COP}^{n^2}$ due to the fact that we can always add nonnegative matrices to any copositive matrix without changing copositivity while we have $D - I \otimes S^* - T^* \otimes I - v^* E_{n^2} \in \mathcal{COP}^{n^2}$ by definition. In fact, it turns out that we have $Q_\varepsilon^* \in \text{int}(\mathcal{COP}^{n^2})$ as for $x \in \mathbb{R}_+^{n^2} \setminus \{0\}$ we have

$$\begin{aligned}
&\mathbf{x}^\mathsf{T}(D - I \otimes S^* - T^* \otimes I - (v^* - \varepsilon)E_{n^2})\mathbf{x} \\
=&\mathbf{x}^\mathsf{T}(D - I \otimes S^* - T^* \otimes I - v^* E_{n^2}))\mathbf{x} + \varepsilon \mathbf{x}^\mathsf{T} E_{n^2}\mathbf{x} \\
\geq&\varepsilon \mathbf{x}^\mathsf{T} E_{n^2}\mathbf{x} > 0.
\end{aligned}$$

From Theorem 1.13 it now follows that there exists an $\bar{r} \in \mathbb{Z}_+$ such that $Q_\varepsilon^* \in \mathcal{C}_{n^2}^{\bar{r}}$.

In order to see that Q_ε^* provides a solution which gives us the optimal value of $\mathcal{GIP}_{COP}$, we consider the objective function of (5.7). First we denote by θ the optimal value of $\mathcal{GIP}_{COP}$, i.e. $\theta = \text{Tr}(S^*) + \text{Tr}(T^*) + n^2 v^*$. Furthermore, let

$$\theta_\varepsilon := \text{Tr}\,(S^*) + \text{Tr}\,(T^*) + n^2(v^* - \varepsilon).$$

Then, because we know from Lemma 5.7 that $\theta \in \mathbb{Z}_+$, we obtain by rounding that

$$\lceil \theta_\varepsilon \rceil = \lceil \text{Tr}\,(S^*) + \text{Tr}\,(T^*) + n^2 v^* - n^2\varepsilon \rceil = \lceil \theta \rceil - \lceil n^2 \varepsilon \rceil = \theta$$

due to the fact that $0 < \varepsilon < \frac{1}{n^2}$. Now, because we also know that $Q_\varepsilon^* \in \mathcal{C}_{n^2}^{\bar{r}}$ it must hold that

$$\lceil \eta^{\bar{r}} \rceil \geq \lceil \theta \rceil.$$

This implies that

$$\theta \geq \lceil \eta^{\bar{r}} \rceil \geq \lceil \theta_\varepsilon \rceil = \theta.$$

Finally, let (S, T, v) be a feasible solution of (5.13) for some $r \in \mathbb{Z}_+$ so that $D - I \otimes S - T \otimes I - vE_{n^2} \in \mathcal{C}_n^r$. Then the polynomial so $\mathbf{x}^\mathsf{T}(D - I \otimes S - T \otimes I - vE_{n^2})\mathbf{x}$ has nonnegative coefficients so that for any $X \in \mathcal{P}_n$ we have

$$0 \le \langle D - I \otimes S - T \otimes I - vE_{n^2}, \mathbf{x}\mathbf{x}^\mathsf{T} \rangle = \langle D, \mathbf{x}\mathbf{x}^\mathsf{T} \rangle - (\mathrm{Tr}(S) + \mathrm{Tr}(T) + vn^2).$$

Then because $\langle D, \mathbf{x}\mathbf{x}^\mathsf{T} \rangle \ge 0$ it follows that $\eta^r \ge \mathrm{Tr}(S) + \mathrm{Tr}(T) + vn^2 \ge 0$ for all $r \in \mathbb{Z}_+$. This implies that we do not need to round η^r to verify isomorphism. This proves our statement. □

In an attempt to obtain an explicit upper bound on $r \in \mathbb{Z}_+$ so that we can solve the $\mathcal{GIP}$ by computing η^r instead of $\mathcal{GIP}_{COP}$ we now give the following lemma.

Lemma 5.17. *Given two graphs $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$, let D and Q_ε^* be the matrices as defined in Proposition 5.16. Then*

$$\kappa := \min_{\mathbf{x} \in \Delta} \mathbf{x}^\mathsf{T} Q_\varepsilon^* \mathbf{x} \ge \frac{1}{n^2}.$$

Proof. We have

$$\begin{aligned}
\kappa &= \min_{\mathbf{x} \in \Delta} \mathbf{x}^\mathsf{T}(D - I \otimes S^* - T^* \otimes I - (v^* - \varepsilon)E_{n^2})\mathbf{x} \\
&= \min_{\mathbf{x} \in \Delta} \mathbf{x}^\mathsf{T}(D - I \otimes S^* - T^* \otimes I - v^* E_{n^2})\mathbf{x} + \varepsilon \mathbf{x}^\mathsf{T} E_{n^2} \mathbf{x} \\
&\ge \min_{\mathbf{x} \in \Delta} \varepsilon \mathbf{x}^\mathsf{T} E_{n^2} \mathbf{x} = \frac{1}{n^2}.
\end{aligned}$$

□

In order to apply Corollary 5.15, we would also need an upper bound on $L = \max_{i,j} |(Q_\varepsilon^*)_{ij}|$. Unfortunately however we have not been able to find such an upper bound. From computational experiments we found that L can be quite large. In fact, it would seem that for the general case L is unbounded as we can always add the positive semidefinite matrix $nI_{n^2} - E_{n^2}$ to Q_ε^* without altering copositivity of Q_ε^*, while simultaneously keeping the objective function of (5.7) unchanged. We expect that is possible to bound the problem in some manner to resolve this problem, but it does not look as if there is a straightforward way of doing so. Note, however, that the bound given in Corollary 5.15 is not tight in general. In fact, in many practical cases a much lower level of the hierarchy is sufficient than that given by Corollary 5.15.

5.5 Reformulating the copositive formulation

In this section we will reformulate the copostive programming problem (5.7) by investigating its dual, that is the completely positive formulation (5.8). Recall that the feasible set of (5.8), which we denoted as $\mathcal{F}$, can be written equivalently as in (5.6). Using this alternative definition for $\mathcal{F}$, we will now directly dualize this set, after which we shall substitute this dual for $\mathcal{COP}^{n^2}$ in our copositive program (5.7). Doing this will allow us to obtain an upper bound for the copositive formulation.

Proposition 5.18. *Let* $\mathcal{F}_n = \operatorname{conv}\left\{\mathbf{y}\mathbf{y}^\mathsf{T} \mid \mathbf{y} = \operatorname{Vec}(X),\, X \in \mathcal{P}_n\right\}$. *Then its dual cone is the set*

$$\mathcal{F}_n^* = \left\{A \in \mathbb{S}^{n^2} \mid \langle A, \mathbf{x}\mathbf{x}^\mathsf{T}\rangle \geq 0 \textit{ for all } \mathbf{x} = \operatorname{Vec}(X),\, X \in \mathcal{P}_n\right\}. \tag{5.16}$$

Proof. From Definition 1.7 we know that the dual of a set $\mathcal{F}$ is defined as

$$\mathcal{F}^* = \left\{A \in \mathbb{S}^{n^2} \mid \langle A, Y\rangle \geq 0 \text{ for all } Y \in \mathcal{F}\right\}.$$

Next, consider the cone

$$\mathcal{H} = \left\{A \in \mathbb{S}^{n^2} \mid \langle A, \mathbf{x}\mathbf{x}^\mathsf{T}\rangle \geq 0, \text{ for all } \mathbf{x} = \operatorname{Vec}(X),\, X \in \mathcal{P}_n\right\}.$$

We will prove our proposition by showing that $\mathcal{F}^* = \mathcal{H}$. First let $A \in \mathcal{H}$. Then for any $Y \in \mathcal{F}$ we can write $Y = \sum_{X \in \mathcal{P}_n} \lambda_X \operatorname{Vec}(X)\operatorname{Vec}(X)^\mathsf{T}$ with $\lambda_X \geq 0$ for all $X \in \mathcal{P}_n$ and $\sum_{X \in \mathcal{P}_n} \lambda_X = 1$. Hence we get

$$\langle A, Y\rangle = \left\langle A, \sum_{X \in \mathcal{P}_n} \lambda_X \operatorname{Vec}(X)\operatorname{Vec}(X)^\mathsf{T}\right\rangle = \sum_{X \in \mathcal{P}_n} \lambda_X \left\langle A, \operatorname{Vec}(X)\operatorname{Vec}(X)^\mathsf{T}\right\rangle \geq 0$$

so that $A \in \mathcal{F}^*$ and consequently $\mathcal{H} \subseteq \mathcal{F}^*$.

Next, suppose that $A \in \mathcal{F}^*$. Then $\langle A, Y\rangle \geq 0$ for all $Y \in \mathcal{F}$. Observe that for $X \in \mathcal{P}_n$ we have $\mathbf{x}\mathbf{x}^\mathsf{T} \in \mathcal{F}$. Hence

$$\langle A, \mathbf{x}\mathbf{x}^\mathsf{T}\rangle \geq 0 \quad \text{for all } X \in \mathcal{P}_n,$$

which implies that $A \in \mathcal{H}$ and therefore $\mathcal{F}^* \subseteq \mathcal{H}$. Consequently, $\mathcal{H} = \mathcal{F}^*$. This completes the proof. $\square$

Note that clearly we have $\mathcal{F} \subset \mathcal{CP}$. So automatically, from duality theory, we get that $\mathcal{COP} \subset \mathcal{F}^*$. Hence, simply replacing $\mathcal{COP}^{n^2}$ by $\mathcal{F}_n^*$ in (5.7) gives us an upper bound for the particular copositive program. In fact, as it turns out, these two conic optimization problems are in some sense equivalent. By equivalent, in this case, we mean that both problems return nonnegative values while the optimal solution is equal to 0 if and only if the two underlying graphs are isomorphic. We formally state this result in the following theorem.

Theorem 5.19 ($\mathcal{GIP}$ over $\mathcal{F}^*$)**.** *Consider any two graphs G_1 and G_2 with adjacency matrices A and B respectively, and let D be defined as in Lemma 5.3. Then consider the following problem*

$$\max_{S,T\in\mathbb{S}^n, v\in\mathbb{R}} \left\{ \mathrm{Tr}(S) + \mathrm{Tr}(T) + vn^2 \mid D - I\otimes S - T\otimes I - vE_{n^2} \in \mathcal{F}^* \right\}, \quad (5.17)$$

which we denote as $\mathcal{GIP}_{\mathcal{P}}$. Then the optimal value of $\mathcal{GIP}_{\mathcal{P}}$ is equal to 0 if and only if G_1 is isomorphic to G_2.

Proof. Due to the fact that $\mathcal{COP}^{n^2} \subseteq \mathcal{F}^*$ the optimal value of $\mathcal{GIP}_{\mathcal{P}}$ is an upper bound for the optimal value of (5.7) and is therefore nonnegative. Now let (S, T, v) be a feasible solution of (5.17) and let $X \in \mathcal{P}_n$. Then

$$0 \le \langle D - I\otimes S - T\otimes I - vE_{n^2}, \mathbf{x}\mathbf{x}^\mathsf{T}\rangle = \langle D, \mathbf{x}\mathbf{x}^\mathsf{T}\rangle - (\mathrm{Tr}(S) + \mathrm{Tr}(T) + vn^2)$$

or equivalently

$$\langle D, \mathbf{x}\mathbf{x}^\mathsf{T}\rangle \ge (\mathrm{Tr}(S) + \mathrm{Tr}(T) + vn^2). \quad (5.18)$$

An important observation is that the expression in the right-hand side of (5.18) (which define the feasible set of (5.17)) is also the objective function of (5.17).

Now assume that G_1 is isomorphic to G_2. From Lemma 5.3 we know that in this case there exists an $X \in \mathcal{P}_n$ for which $\langle D, \mathbf{x}\mathbf{x}^\mathsf{T}\rangle = 0$ so that automatically from (5.18) we know that the optimal value of $\mathcal{GIP}_{\mathcal{P}}$ is at most 0. Then, because $S = T = 0$, $v = 0$ defines a feasible solution to (5.17) this must in fact be an optimal solution to (5.17).

Next, assume that G_1 and G_2 are not isomorphic. Then by Lemma 5.3 we know that $\langle D, \mathbf{x}\mathbf{x}^\mathsf{T}\rangle > 0$ for every $X \in \mathcal{P}_n$. Then $S = T = 0$ and $v = \frac{1}{n^2}\min_{X\in\mathcal{P}_n}\langle D, \mathbf{x}\mathbf{x}^\mathsf{T}\rangle$ is a feasible solution of (5.17) with positive objective value, and hence the optimal value of $\mathcal{GIP}_{\mathcal{P}}$ is strictly positive. This finishes the proof. □

Observe that in the proof of Theorem 5.19 explicit feasible solutions are given both when G_1 and G_2 are isomorphic as well when they are non-isomorphic.

Moreover, these solutions are both nonnegative irrespective of D and $X \in \mathcal{P}_n$ so that w.l.o.g. we can assume for (5.17) that $S, T \geq 0$ and $v \geq 0$. Finally, because the right-hand side of the inequality (5.18) is independent of both $X \in \mathcal{P}_n$ as well as D we can simply substitute $\mathrm{Tr}(S) + \mathrm{Tr}(T) + vn^2$ by a single nonnegative variable γ, i.e. both ($\langle I \otimes S - T \otimes I - vE_{n^2}, \mathbf{x}\mathbf{x}^\mathsf{T} \rangle$) in the matrix constraint as well as the objective of (5.17).

More importantly, because $\mathcal{COP} \subset \mathcal{F}^*$, Theorem 5.19 implies that we can substitute the copositive cone in (5.7) by any cone $\mathcal{B} \subseteq \mathbb{R}^{n^2 \times n^2}$ that satisfies $\mathcal{COP} \subseteq \mathcal{B} \subset \mathcal{F}^*$. In particular if there exists such a cone that, moreover is tractable, then $\mathcal{GIP}$ is in P. This presents us with the following interesting open problem.

Open Problem 5.20. *Let*

$$\mathcal{F}_n^* = \left\{ A \in \mathbb{S}^{n^2} \mid \langle A, \mathbf{x}\mathbf{x}^\mathsf{T} \rangle \geq 0, \textit{ for all } \mathbf{x} = \mathrm{Vec}(X),\ X \in \mathcal{P}_n \right\}.$$

Does there exists a computationaly tractable cone $\mathcal{B}_n \subseteq \mathbb{R}^{n^2 \times n^2}$ such that $\mathcal{COP}^{n^2} \subseteq \mathcal{B}_n \subset \mathcal{F}_n^*$? If so, this would imply that $\mathcal{GIP}$ is solvable in polynomial time.

Note that the inequalities $\langle A, \mathbf{x}\mathbf{x} \rangle \geq 0$, $X \in \mathcal{P}_n$ define supporting hyperplanes for the copositive cone. In other words $\mathcal{COP}^{n^2}$ and $\mathcal{F}_n^*$ share parts of their boundaries. In particular, their boundaries touch at an exponential number (in n) of points. This seems to suggest that it will be difficult to create a cone $\mathcal{B}_n$ as suggested in Open Problem 5.20, particularly one for which verifying membership is tractable. However, from Proposition 5.16 we know that there exist finite levels $\bar{r} \in \mathbb{Z}_+$ for each of the hierarchies $\mathcal{C}_n^r$, $\mathcal{Q}_n^r$, and $\mathcal{K}_n^r$, that can be used to construct approximations to the copostive formulation (5.7) that are in fact sufficient to decide the graph isomorphism problem. Hence a sufficient certificate for $\mathcal{GIP}$ to be in Pwould be a tractable cone $\mathcal{B}_n$ such that, say $\mathcal{C}_{n^2}^{\bar{r}} \subseteq \mathcal{B}_n \subseteq \mathcal{F}_n^*$. Whether such a construction is possible seems to depend largely on the upper bound one can obtain for $\bar{r}$ and the geometry of the cones $\mathcal{C}_n^r$, $\mathcal{Q}_n^r$, and $\mathcal{K}_n^r$.

Summary

In this thesis we investigated a number of properties regarding the copositive and completely positive cone, motivated by results obtained in copositive optimization. These two cones are respectively defined as:

$$\mathcal{COP}^n := \{A \in \mathbb{S}^n \mid \mathbf{x}^\mathsf{T} A \mathbf{x} \geq 0 \text{ for all } \mathbf{x} \in \mathbb{R}^n_+\},$$
$$\mathcal{CP}^n := \{A \in \mathbb{S}^n \mid \; A = \sum_{i=1}^{k} \mathbf{b}_i \mathbf{b}_i^\mathsf{T},\, \mathbf{b}_i \geq 0 \text{ for all } i\},$$

and are mutually dual.

We first studied the complexity of the membership problem for both of these cones. It is a known result that verifying copositivity is co-NP-complete. On the other hand, the complexity of verifying that a matrix is completely positive was still unknown. In this thesis we confirm the expected result that the membership problem for the completely positive cone is NP-hard. In fact, we show that both the weak and the strong membership problem for the copositive as well as the completely positive cone belong to the class NP-hard.

We also investigated the property of irreducibility with respect to the cone of nonnegative matrices, a weaker version of extremality. In particular, we provide a necessary and sufficient condition for a copositive matrix to be irreducible. For the 5×5 copositive cone we give a complete characterization of all irreducible matrices. Then, we show that every 5×5 copositive matrix that is not the sum of a nonnegative and a positive semidefinite matrix can be expressed as the sum of a nonnegative and a single irreducible matrix. This latter result is used to show that the 5×5 copositive cone reduces to the level one Parrilo cone under specific scalings.

We furthermore proved the result that we can scale any matrix, that is not the sum of a nonnegative and a positive semidefinite matrix, out of any level r of the Parrilo cone for $r \geq 1$ and $n \geq 5$. For the level one Parrilo cone an explicit way to construct such scalings is provided. We then investigated scalings in the opposite direction and find that an algorithm that constructs such scalings cannot have a polynomial running time unless P= NP. We introduce the concept of non-decreasing scalings, which are scalings that cannot scale a matrix out of any level of a hierarchy it is in. We moreover give an example of a non-decreasing scaling that is capable of scaling matrices downward in (amongst others) the Parrilo hierarchy.

Finally, we provide an application in the form of a copositive formulation of the graph isomorphism problem and show that we can in fact decide isomorphism using finite levels of some approximation hierarchies of the copositive cone. Then, several alternative formulations are suggested, one of which implies a potential method to construct a certificate that the graph isomorphism problem is in P.

Samenvatting

In deze thesis onderzoeken wij een aantal eigenschappen met betrekking tot de copositieve en compleet positieve kegel, gemotiveerd door resultaten binnen de copositieve optimalisatie. Deze twee kegels zijn respectievelijk gedefinieerd als:

$$\begin{aligned} \mathcal{COP}^n &:= \{A \in \mathbb{S}^n \mid \mathbf{x}^\mathsf{T} A\mathbf{x} \geq 0 \text{ voor iedere } \mathbf{x} \in \mathbb{R}^n_+\}, \\ \mathcal{CP}^n &:= \{A \in \mathbb{S}^n \mid \ A = \sum_{i=1}^{k} \mathbf{b}_i\mathbf{b}_i^\mathsf{T},\, \mathbf{b}_i \geq 0 \text{ voor iedere } i\}, \end{aligned}$$

en zijn elkaars duale kegel.

We bestudeerden als eerste de complexiteit van het lidmaatschap probleem van de beide kegels. Het is een bekend resultaat dat de complexiteit van het verifiëren van copositiviteit co-NP-compleet is. Daarentegen was de complexiteit van het nagaan of een matrix compleet positief is onbekend. In deze thesis bevestigen we het verwachte resultaat dat het lidmaatschap probleem van de compleet positieve kegel NP-moeilijk is. Sterker, we laten zien dat zowel het zwakke als het sterke lidmaatschap probleem voor zowel de copositieve alsmede de compleet positieve kegel tot de klasse NP-moeilijk behoort.

We onderzochten tevens de eigenschap van onverkleinbaarheid met betrekking tot de kegel van niet-negatieve matrices, zijnde een zwakkere versie van extremaliteit. In het bijzonder verstrekken wij een voldoende en noodzakelijke voorwaarde voor een copositieve matrix om onverkleinbaar te zijn. Voor de 5×5 copositieve kegel geven wij een volledige beschrijving van alle onverkleinbare matrices. Daarnaast, laten wij zien dat elke 5×5 copositieve matrix die niet de som is van een niet-negatieve en een positief-semidefiniete matrix kan worden geschreven als de som van een niet-negatieve matrix en een enkele onverkleinbare matrix. Het laatstgenoemde resultaat hebben wij vervolgens gebruikt om te laten zien dat de copositieve kegel kan worden gereduceerd tot de niveau één Parrilo kegel met behulp van specifieke schalingen.

We bewijzen verder het resultaat dat we elke matrix, welke niet de som van een niet-negatieve en een positief-semidefiniete matrix is, uit elke willekeurige niveau r Parrilo kegel kan worden geschaald voor $r \geq 1$ en $n \geq 5$. Voor de niveau één Parrilo kegel geven we een expliciete manier om dergelijke schalingen te kunnen construeren. Vervolgens onderzochten we schalingen in de tegenovergestelde richting en leerden we dat een algoritme dat dergelijke schalingen kan produceren geen polynomiale rekentijd kan hebben tenzij P= NP. We introduceerde het begrip van niet-negatieve schalingen, dit zijn schalingen die niet in staat zijn een matrix uit een bepaald niveau van een hiërarchie te schalen. We geven daarnaast een voorbeeld van een niet-negatieve schaling die in staat is om matrices naar beneden te schalen in (onder andere) de Parrilo hiërarchie.

Tenslotte verstrekken we een toepassing in de vorm van een copositieve formulering van het graaf isomorphisme probleem en laten we zien dat we isomorphisme kunnen vaststellen met behulp van een eindig niveau van enkele benadering hiërarchieën voor de copositieve kegel. Vervolgens worden er en aantal alternatieve formuleringen voorgesteld, waarvan er één een mogelijke methode impliceert om een certificaat te construeren dat laat zien dat het graaf isomorphisme probleem tot P behoort.

Bibliography

[ACE95] Lars-Erik Andersson, Gengzhe Chang, and Tommy Elfving. Criteria for copositive matrices using simplices and barycentric coordinates. *Linear Algebra and its Applications*, 220:9–30, 1995.

[AK06] Vikraman Arvind and Piyush P. Kurur. Graph isomorphism is in SPP. *Information and Computation*, 204(5):835–852, 2006.

[AM14] Amir Ali Ahmadi and Anirudha Majumdar. DSOS and SDSOS optimization: LP and SOCP-based alternatives to sum of squares optimization. *48th Annual Conference on Information Sciences and Systems (CISS)*, pages 1–5, 2014.

[BAD09] Samuel A. Burer, Kurt M. Anstreicher, and Mirjam Dür. The difference between 5×5 doubly nonnegative and completely positive matrices. *Linear Algebra and its Applications*, 431:1539–1552, 2009.

[Bas69] Victor J.D. Baston. Extreme copositive quadratic forms. *Acta Arithmetica*, XV:319–327, 1969.

[Bau65] Leonard D. Baumert. *Extreme Copositive Quadratic Forms*. PhD thesis, California Institute of Technology, Pasadena, California, 1965.

[Bau66] Leonard D. Baumert. Extreme copositive quadratic forms. *Pacific Journal of Mathematics*, 19(2):197–204, 1966.

[Bau67] Leonard D. Baumert. Extreme copositive quadratic forms. II. *Pacific Journal of Mathematics*, 20(1):1–20, 1967.

[BD08] Stefan Bundfuss and Mirjam Dür. Algorithmic copositivity detection by simplicial partition. *Linear Algebra and its Applications*, 428:1511–1523, 2008.

[BD09] Stefan Bundfuss and Mirjam Dür. An adaptive linear approximation algorithm for copositive programs. *SIAM Journal on Optimization*, 20(1):30–53, 2009.

[BDdK$^+$00] Immanuel M. Bomze, Mirjam Dür, Etienne de Klerk, Cornelis Roos, Arie J. Quist, and Tamás Terlaky. On copositive programming and standard quadratic optimization problems. *Journal of Global Optimization*, 18:301–320, 2000.

[BdK02] Immanuel M. Bomze and Etienne de Klerk. Solving standard quadratic optimization problems via linear, semidefinite and copositive programming. *Journal of Global Optimization*, 24(2):163–185, 2002.

[BE12] Immanuel M. Bomze and Gabriele Eichfelder. Copositivity detection by difference-of-convex decomposition and ω-subdivision. *Mathematical Programming*, 138:365–400, 2012.

[Ber09] Dennis S. Bernstein. *Matrix Mathematics: Theory, Facts and Formulas (second edition)*. Princeton University Press, 2009.

[BGM82] László Babai, D. Yu. Grigoryev, and David M. Mount. Isomorphism of graphs with bounded eigenvalue multiplicity. In *Proceedings of the Fourteenth Annual ACM Symposium on Theory of Computing*, STOC '82, pages 310–324, New York, NY, USA, 1982. ACM.

[BJ10] Immanuel M. Bomze and Florian Jarre. A note on Burer's copositive representation of mixed-binary QPs. *Optimization Letters*, 4(3):465–472, 2010.

[BJR11] Immanuel M. Bomze, Florian Jarre, and Franz Rendl. Quadratic factorization heuristics for copositive programming. *Mathematical Programming Computation*, 3(1):37–57, 2011.

[BM76] John A. Bondy and Uppaluri S.R. Murty. *Graph Theory with Applications*. The Macmillan Press Ltd., 1976.

[Bom87] Immanuel M. Bomze. Remarks on the recursive structure of copositivity. *Journal of Information and Optimization Sciences*, 8(3):243–260, 1987.

[Bom89] Immanuel M. Bomze. Copositivity and optimization. *Methods in Operations Research*, 58:27–35, 1989.

[Bom12] Immanuel M. Bomze. Copositive programming - recent developments and applications. *European Journal of Operations Research*, 216:509–520, 2012.

[BSM03] Abraham Berman and Naomi Shaked-Monderer. *Completely Positive Matrices*. World Scientific Pub. Co. Pte. Inc., 2003.

[BSU12] Immanuel M. Bomze, Werner Schachinger, and Gabriele Uchida. Think co(mpletely)positive ! matrix properties, examples and a clustered bibliography on copositive optimization. *Journal of Global Optimization*, 52(3):423–445, 2012.

[BSU14a] Immanuel M. Bomze, Werner Schachinger, and Reinhard Ullrich. From seven to eleven: completely positive matrices with high CP-rank. *Linear Algebra and its Applications*, 459(15):208–221, 2014.

[BSU14b] Immanuel M. Bomze, Werner Schachinger, and Reinhard Ullrich. New lower bounds and asymptotics for the CP-rank. `http://www.newton.ac.uk/preprints/NI14048.pdf`, 2014. Preprint.

[Bur09] Samuel A. Burer. On the copositive representation of binary and continuous nonconvex quadratic programs. *Mathematical Programming*, 120:479–495, 2009.

[Bur12] Samuel A. Burer. Copositive programming. In M.F. Anjos and J.B. Lasserre, editors, *Handbook of Semidefinite, Conic and Polynomial Optimization: Theory, Algorithms, Software and Applications*, pages 201–218. Springer, New York, 2012.

[CFI92] Jin-Yi Cai, Martin Fürer, and Neil Immerman. An optimal lower bound on the number of variables for graph identification. *Combinatorica*, 12(4):389–410, 1992.

[CHL70] Richard W. Cottle, G.J. Habetler, and C.E. Lemke. On classes of copositive matrices. *Linear Algebra and its Applications*, 3:295–310, 1970.

[CS94] Gengzhe Chang and Thomas W. Sederberg. Nonnegative quadratic Bézier triangular patches. *Computer Aided Geometric Design*, 11:113–116, 1994.

[DA13] Hongbo Dong and Kurt Anstreicher. Separating doubly nonnegative and completely positive matrices. *Mathematical Programming*, 137(1):131–153, 2013.

[DD12] Peter J.C. Dickinson and Mirjam Dür. Linear-time complete positivity detection and decomposition of sparse matrices. *SIAM Journal on Matrix Analysis and Applications*, 33(3):701–720, 2012.

[DDGH13a] Peter J.C. Dickinson, Mirjam Dür, Luuk Gijben, and Roland Hildebrand. Irreducible elements of the copositive cone. *Linear Algebra and its Applications*, 439(6):1605–1626, 2013.

[DDGH13b] Peter J.C. Dickinson, Mirjam Dür, Luuk Gijben, and Roland Hildebrand. Scaling relationship between the copositive cone and Parrilo's first level approximation. *Optimization Letters*, 7(8):1669–1679, 2013.

[DDV14] Christian Dobre, Mirjam Dür, and Frank Vallentin. A copositive formulation for the stability number of infinite graphs. `http://arxiv.org/pdf/1305.1819v1.pdf`, 2014. Preprint.

[DG14] Peter J.C. Dickinson and Luuk Gijben. On the computational complexity of membership problems for the completely positive cone and its dual. *Computational Optimization and Applications*, 57(2):403–415, 2014.

[Dia62] Palahenedi H. Diananda. On nonnegative forma in real variables some or all of which are nonnegative. *Mathematical Proceedings of the Cambridge Philosophical Society*, 58(1):17–25, 1962.

[Dic10] Peter J. C. Dickinson. An improved characterisation of the interior of the completely positive cone. *Electronic Journal of Linear Algebra*, 20:723–729, 2010.

[Dic11] Peter J.C. Dickinson. Geometry of the copositive and completely positive cone. *Journal of Mathematical Analysis and Applications*, 380(1):377–395, 2011.

[Dic13] Peter J.C. Dickinson. *The Copositive Cone, the Completely Positive Cone and their Generalisations*. PhD thesis, University of Groningen, 2013.

[DJL94] John H Drew, Charles R. Johnson, and Raphael Loewy. Completely positive matrices associated with M-matrices. *Linear and Multilinear Algebra*, 37(4):303–310, 1994.

[dK08] Etienne de Klerk. The complexity of optimizing over a simplex, hypercube or sphere: a short survey. *Central European Journal of Operations Research*, 16(2):111–125, 2008.

[dKP02] Etienne de Klerk and Dmitrii V. Pasechnik. Approximation of the stability number of a graph via copositive programming. *SIAM Journal on Optimization*, 12(4):875–892, 2002.

[DLC14] Bo Dong, Matthew M. Lin, and Moody T. Chu. Nonnegative rank factorization - a heuristic approach via rank reduction. *Numerical Algorithms*, 65(2):251–274, 2014.

[Don13] Hongbo Dong. Symmetric tensor approximation hierarchies for the completely positive cone. *SIAM Journal on Optimization*, 23(3):1850–1866, 2013.

[DR10] Igor Dukanovic and Franz Rendl. Copositive programming motivated bounds on the stability and the chromatic numbers. *Mathematicl Programming*, 121(2):249–268, 2010.

[DS08] Mirjam Dür and Georg Still. Interior points of the completely positive cone. *Electronic Journal of Linear Algebra*, 17:48–53, 2008.

[Dür10] Mirjam Dür. Copositive programming - a survey. In M. Diehl, F. Glineur, E. Jarlebring, and W. Michiels, editors, *Recent Advances in Optimization and its Applications in Engineering*, pages 3–20. Springer-Verlag Berlin Heidelberg, 2010.

[DV13] Christian Dobre and Juan Vera. Exploiting symmetry in copositive programs via semidefinite hierarchies. `http://www.optimization-online.org/DB_FILE/2013/12/4181.pdf`, 2013. Preprint.

[EJ08] Gabriele Eichfelder and Johannes Jahn. Set-semidefinite optimization. *Journal of Convex Analysis*, 15:767–801, 2008.

[Gad58] Jerry W. Gaddum. Linear inequalities and quadratic forms. *Pacific Journal of Mathematics*, 8(3):411–414, 1958.

[GJ79] Michael R. Garey and David S. Johnson. *Computers and Intractability: A Guide to the Theory of $\mathcal{NP}$-Completeness*. W.H. Freeman and co., 1979.

[GL07] Nebojsa Gvozdenović and Monique Laurent. Semidefinite bounds for the stability number of a graph via sums of squares of polynomials. *Mathematical Programming*, 110(1):145–173, 2007.

[GL08] Nebojsa Gvozdenović and Monique Laurent. The operator Ψ for the chromatic number of a graph. *SIAM Journal on Optimization*, 19(2):572–591, 2008.

[GLS88] Martin Grötschel, Lászlo Lovász, and Alexander Schrijver. *Geometric Algorithms and Combinatorial Optimization*. Springer-Verlag (Berlin), 1988.

[HH69] Emilie Haynsworth and Alan J. Hoffman. Two remarks on copositive matrices. *Linear Algebra and its Applications*, 2:387–392, 1969.

[Hil12] Roland Hildebrand. The extreme rays of the 5×5 copositive cone. *Linear Algebra and its Applications*, 437(7):1538–1547, 2012.

[HJ62] Marshall Hall Jr. Discrete problems. In *Survey of Numerical Analysis John Todd (Ed.)*, pages 518–542. McGraw-Hill Book Company Inc., 1962.

[HJR05] Leslie Hogben, Charles R. Johnson, and Robert Reams. The copositive completion problem. *Linear Algebra and its Applications*, 408:207–211, Oct 2005.

[HN63] Marshall Hall, Jr. and Morris Newman. Copositive and completely positive quadratic forms. *Mathematical Proceedings of the Cambridge Philosophical Society*, 59:329–339, 1963.

[Hor76] Reiner Horst. An algorithm for nonconvex programming problems. *Mathematical Programming*, 10(1):312–321, 1976.

[Hor97] Reiner Horst. On generalized bisection of n-simplices. *Mathematics of Computation*, 66(218):691–698, 1997.

[HPT95] Reiner Horst, Panos M. Pardalos, and Nguyen Van Thoai. *Introduction to Global Optimization*. Nonconvex Optimization and Its Applications Vol. 3. Springer, 1995.

[HRMP95] Christoph Helmberg, Franz Rendl, Bojan Mohar, and Svatopluk Poljak. A spectral approach to bandwidth and separator problems in graphs. *Linear and Multilinear Algebra*, 39:73–90, 1995.

[HT71] John E. Hopcroft and Robert E. Tarjan. A v^2 algorithm for determining isomorphism of planar graphs. *Information Processing Letters*, 1(1):32–34, 1971.

[HUS10] Jean-Baptiste Hiriart-Urruty and Alberto Seeger. A variational approach to copositive matrices. *SIAM Review*, 52(4):593–629, 2010.

[HW74] John E. Hopcroft and J. K. Wong. Linear time algorithm for isomorphism of planar graphs (preliminary report). In *Proceedings of the Sixth Annual ACM Symposium on Theory of Computing*, STOC '74, pages 172–184, New York, NY, USA, 1974. ACM.

[JR08] Charles R. Johnson and Robert Reams. Constructing copositive matrices from interior matrices. *Electronic Journal of Linear Algebra*, 17:9–20, 2008.

[JS09] Florian Jarre and Katrin Schmallowsky. On the computation of $\mathcal{C}^*$ certificates. *Journal of Global Optimization*, 45(2):281–296, 2009.

[Kay87] Mohammad Kaykobad. On nonnegative factorization of matrices. *Linear Algebra and its Applications*, 96:27–33, 1987.

[Kea78] Baker Kearfott. A proof of convergence and an error bound for the method of bisection in $\mathbb{R}^n$. *Mathematics of Computation*, 32(144):1147–1153, 1978.

[KG12] Vassilis Kalofolias and Efstratios Gallopoulos. Computing symmetric nonnegative rank factorizations. *Linear Algebra and its Applications*, 436:421–435, 2012.

[Kha79] Leonid Khachiyan. A polynomial algorithm in linear programming (in russian). *Doklady Akademii Nauk SSSR*, 244:1093–1096, 1979.

[Las11] Jean-Bernard Lasserre. A new look at nonnegativity on closed sets and polynomial optimization. *SIAM Journal on Optimization*, 21(3):864–885, 2011.

[Las14] Jean-Bernard Lasserre. New approximations for the cone of copositive matrices and its dual. *Mathematical Programming*, 144(1):265–276, 2014.

[Lau08] Monique Laurent. Semidefinite programming in combinatorial and polynomial optimization. *Nieuw Archief voor Wiskunde (NAW)*, 5(4), 2008.

[Lau09] Monique Laurent. Sums of squares, moment matrices and optimization over polynomials. In *Emerging applications of algebraic geometry*, volume 149 of *IMA Vol. Math. Appl.*, pages 157–270. Springer, New York, 2009.

[Luk82] Eugene M. Luks. Isomorphism of graphs of bounded valence can be tested in polynomial time. *Journal of Computer and System Sciences*, 25(1):42–65, 1982.

[MA13] Shashank K. Mehta and Pawan Aurora. Completely positive formulation of the graph isomorphism problem. `http://arxiv.org/pdf/1301.2390v1.pdf`, 2013. Preprint.

[McK81] Brendan D. McKay. Practical graph isomorphism. *Congressus Numerantium*, 30:45–87, 1981.

[Miy95] Takunari Miyazaki. The complexity of McKay's canonical labeling algorithm. In *Groups and Computation II*, volume 28 of *DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, pages 239–256. American Mathematical Society, 1995.

[MK87] Katta G. Murty and Santosh N. Kabadi. Some $\mathcal{NP}$-complete problems in quadratic and nonlinear programming. *Mathematical Programming*, 39(2):117–129, 1987.

[Mot52] Theodore S. Motzkin. Copositive quadratic forms. *National Bureau of Standards Report*, 1818:11–22, 1952.

[NTZ11] Karthik Natarajan, Chung-Piaw Teo, and Zhichao Zheng. Mixed 0-1 linear programs under objective uncertainty: a completely positive representation. *Operations Research*, 59:713–728, 2011.

[Par00] Pablo A. Parrilo. *Structured Semidefinite Programs and Semialgebraic Geometry Methods in Robustness and Optimization.* PhD thesis, California Institute of Technology, 2000.

[Pól28] György Pólya. Über Positive Darstellung von Polynomen. *Vierteljschr. Naturforsch. Ges. Zürich*, 73:141–145, 1928. Collected Papers, Vol. 2, MIT Press, Cambridge, MA, London, 1974, pp. 309-313.

[PR01] Victoria Powers and Bruce Reznick. A new bound for Pólya's theorem with applications to polynomials positive on polyhedra. *Journal of Pure and Applied Algebra*, 164:221–229, 2001.

[PR07] Janez Povh and Franz Rendl. A copositive programming approach to graph partitioning. *SIAM Journal on Optimization*, 18(1):223–241, 2007.

[PR09] Janez Povh and Franz Rendl. Copositive and semidefinite relaxations of the quadratic assignment problem. *Discrete Optimization*, 6:231–241, 2009.

[PVZ06] Javier Peña, Juan Vera, and Luis F. Zuluaga. LMI approximations for cones of positive semidefinite forms. *SIAM Journal on Optimization*, 16:1076–1091, 2006.

[PVZ07] Javier Peña, Juan Vera, and Luis F. Zuluaga. Computing the stability number of a graph via linear and semidefinite programming. *SIAM Journal on Optimization*, 18(1):87–105, 2007.

[PY93] Li Ping and Feng Yu Yu. Criteria for copositive matrices of order four. *Linear Algebra and its Applications*, 194:109–124, 1993.

[Rob73] Raphael M. Robinson. Some definite polynomials which are not sums of squares of real polynomials. In *Selected questions of algebra and logic (collection dedicated to the memory of A. I. Mal′cev) (Russian)*, pages 264–282. Izdat. "Nauka" Sibirsk. Otdel., Novosibirsk, 1973.

[Sch88] Uwe Schöning. Graph isomorphism is in the low hierarchy. *Journal of Computer amd System Sciences*, 37(3):312–323, 1988.

[Sch03] Alexander Schrijver. *Combinatorial optimization: polyhedra and efficiency*, volume 24 of *Algorithms and Combinatorics*. Springer, 2003.

[Sch09] Pascal Schweitzer. *Problems of Unknown Complexity, Graph Isomorphism and Ramsey Theoretic Numbers*. PhD thesis, Universität des Saarlandes, 2009.

[Sho77] Naum Z. Shor. Cut-off method with space extension in convex programming problems. *Cybernetics*, 13:94–96, 1977.

[SM09] Naomi Shaked-Monderer. A note on upper bounds on the CP-rank. *Linear Algebra and its Applications*, 431:2407–2413, 2009.

[SMBJS13] Naomi Shaked-Monderer, Immanuel M. Bomze, Florian Jarre, and Werner Schachinger. On the CP-rank and the minimal cp factorizations of a completely positive matrix. *SIAM Journal on Matrix Analysis and Applications*, 34(2):355–368, 2013.

[TH88] Hoang Tuy and Reiner Horst. Convergence and restart in branch-and-bound algorithms for global optimization. Application to concave minimization and d.c. optimization problems. *Mathematical Programming*, 41(1):161–183, 1988.

[Tuy88] Hoang Tuy. Effect of the subdivision strategy on convergence and efficiency of some global optimization algorithms. *Journal of Global Optimization*, 1(1):23–36, 1988.

[WL68] Boris Weisfeiler and A. A. Lehman. A reduction of a graph to a canonical form and an algebra arising during this reduction (in Russian). *Nauchno-Technicheskaya Informatsia*, 2(9):12–16, 1968.

[Yca82] Bernard Ycart. Extrémales du cône des matrices de type non négatif, à coefficients positifs ou nuls. *Linear Algebra and its Applications*, 48:317–330, 1982.

[YN76] D.B. Yudin and A.S. Nemirovskii. Informational complexity and efficient methods for the solution of convex extremal problems. *Ekonomika i Matematicheskie Metody*, 12(2):357–369, 1976. English translation: Matekon 13 (3) (1977) 25-45.

Nomenclature

Functions and Operators

$\mathrm{Aut}(\mathbb{R}^n_+)$	The automorphism group of the nonnegative orthant, page 40
δ_{ij}	The Kronecker delta, page 86
Diag	Operator that turns a vector $\mathbf{d}$ into a diagonal matrix D such that $D_{ii} = d_i$ for all i, page 3
diag	Operator that creates a vector $\mathbf{a}$ from a matrix A such that $a_i = A_{ii}$ for all i, page 3
$\langle A, B \rangle$	Inner product of A and B, page 3
$\mathbf{x}^{\boldsymbol{\beta}}$	For $\mathbf{x} \in \mathbb{R}^n$ and $\boldsymbol{\beta} \in \mathbb{Z}^n_+$, we have $\mathbf{x}^{\boldsymbol{\beta}} = \prod_{i=1}^n x_i^{\beta_i}$, page 16
$\mathcal{O}(\bullet)$	The big-O notation describing limit behavior of functions, page 3
$\mathbb{R}[\mathbf{x}]$	Ring of polynomials in $\mathbf{x}$ with coefficients in $\mathbb{R}$, page 3
$\mathrm{Tr}(A)$	Trace of A, page 3
Vec	operator that creates a vector from a matrix by stacking its columns, page 3
$\lvert\mathbf{a}\rvert$	Sum of the absolute values of the elements of $\mathbf{a}$, page 3
$A \circ B$	Hadamard product of the matrices A and B, page 3
$A \otimes B$	Kronecker product of the matrices A and B, page 3
$r^*_{\mathbb{Y}^r_n}(A)$	For a hierarchy $\mathbb{Y}^r_n$ approximating the copositive cone, $r^*_{\mathbb{Y}^r_n}(A)$ is the smallest $\bar{r}$ such that $A \in \mathbb{Y}^{\bar{r}}_n$, page 20
$\mathcal{G}_n$	Group of automorphisms w.r.t. scalings and permutations, page 40
CP-rank(A)	The CP-rank of a matrix A, page 8

SOS	sum of squares, page 15

Graph notation

α_G	The stability number of a graph G, page 23
$\bar{G} = (V, \bar{E})$	The complement graph of $G = (V, E)$, page 3
A_G	The adjacency matrix of a graph G, page 23
$G = (V, E)$	Graph with vertex set V and edge set E, page 3

Matrix notation

$(\mathbb{Y}_n^r)_{r\in\mathbb{Z}_+}$	A placeholder for any hierarchy of cones approximating the copositive cone from within, page 20
$\mathcal{CP}^n$	The completely positive cone of order n, page 8
$\mathcal{COP}^n$	The copositive cone of order n, page 4
$\mathcal{C}_n^r$	Approximation hierarchy for the copositive cone, page 14
$\mathbf{A}_{\bullet i}$	The i^{th} column of the matrix A, page 86
$\mathbf{A}_{i\bullet}$	The i^{th} row of the matrix A, page 86
$\mathcal{D}$	The set of $n \times n$ positive diagonal matrices, page 67
$\widetilde{\mathcal{N}}^n$	Set of $n \times n$ nonnegative real matrices with zero diagonal, page 41
$\mathcal{N}^n$	Set of $n \times n$ nonnegative real matrices, page 2
$\mathcal{K}_n^r$	The Parrilo r cone of order n, page 15
$\mathcal{P}_n$	the set of $n \times n$ permutation matrices, page 3
$\mathcal{S}_+^n$	The semidefinite cone of order n, page 4
$\mathbb{R}^{n\times m}$	Set of real $n \times m$ matrices, page 2
$\mathbb{S}^n$	Set of $n \times n$ symmetric real matrices, page 2
$\mathcal{Q}_n^r$	Approximation hierarchy for the copositive cone, page 16
$A_\mathcal{I}$	Principle submatrix of A containing elements whose row and column indices are in the set $\mathcal{I}$, page 40

E_n^{ij} The all zero matrix of order n except for the $(i,j)^{th}$ position which is equal to 1, page 25

$E_{\{i,j\}}$ The all zero matrix except for the $(i,j)^{th}$ and $(j,i)^{th}$ position, page 40

E_n n-dimensional all ones matrix, page 3

I_n n-dimensional identity matrix, page 3

$S(\boldsymbol{\theta})$ The Hildebrand matrices, page 46

dd diagonally dominant, page 9

H The Horn matrix, page 12

LMI Linear matrix inequality, page 16

sdd scaled diagonally dominant, page 9

Set notation

$\mathrm{cl}(K)$ Closure of the set K, page 3

$\mathrm{conv}(K)$ Convex hull of the set K, page 3

$\mathrm{cone}(K)$ Conic hull of the set K, page 3

$\dim(K)$ Dimension of the set K, page 3

$\mathrm{Ext}(K)$ The set of extreme rays of a set K, page 11

$\mathrm{int}(K)$ Interior of the set K, page 3

$\mathcal{V}^A$ Set of zeros of a matrix A, page 40

$\mathrm{supp}(\mathcal{M})$ The support of a set $\mathcal{M}$, page 40

Vector notation

$(\mathbf{x}^{[\leq d]})$ Vector containing all monomials of degree at most d, page 15

$(\mathbf{x}^{[d]})$ Vector containing all monomials of degree exactly d, page 17

$\mathrm{S}\,(K,-\varepsilon)$ ε inner approximation of the set K for some $\varepsilon \in \mathbb{R}_{++}$, page 28

$\mathrm{S}\,(K,\varepsilon)$ ε outer approximation of the set K for some $\varepsilon \in \mathbb{R}_{++}$, page 28

Δ_n The n dimensional standard simplex, page 2

$\mathbb{Z}^n$	Set of integer n-vectors, page 2
$\mathbf{e}_n$	n-dimensional vector of all ones, page 3
$\mathcal{Q}$	Either $\mathbb{Q}^n$ or $(\mathbb{Q}^{n\times n} \cap \mathbb{S}^n)$, page 28
$\mathbb{Q}^n$	Set of rational n-vectors, page 2
$\mathbb{R}^n$	Set of real n-vectors, page 2
$\mathbb{R}^n_+$	Set of nonnegative real n-vectors, page 2
$\mathbb{R}^n_{++}$	Set of positive real n-vectors, page 2
$\mathrm{supp}(\boldsymbol{u})$	The support of $\boldsymbol{u}$, page 40

Index